Peptidic inhibitors of therapeutically relevant proteases: Design, synthesis, and functional evaluation

TECHNISCHE UNIVERSITÄT DARMSTADT

Vom Fachbereich Chemie
Der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines
Doctor rerum naturalium (Dr. rer. nat)

genehmigte
kumulative Dissertation

vorgelegt von

M. Sc. Heiko Fittler
aus Aschaffenburg

Referent:	Prof. Dr. Harald Kolmar
Korreferent:	Prof. Dr. Katja Schmitz

Tag der Einreichung:	10. Dezember 2015
Tag der mündlichen Prüfung:	8. Februar 2016

Darmstadt 2016

D17

Die vorliegende Arbeit wurde unter der Leitung von Herrn Prof. Dr. Harald Kolmar am Clemens Schöpf-Institut für Organische Chemie und Biochemie der Technischen Universität Darmstadt von April 2012 bis Dezember 2015 angefertigt.

Teile der vorliegenden Arbeit sind bereits veröffentlicht

Bernhard Valldorf, **Heiko Fittler**, Lukas Deweid, Aileen Ebenig, Stephan Dickgießer, Carolin Sellmann, Janine Becker, Stefan Zielonka, Martin Empting, Olga Avrutina, and Harald Kolmar, An Apoptosis Inducing Peptidic Heptad that Efficiently Clusters Death Receptor 5, *Angewandte Chemie* **2016**, accepted.

Heiko Fittler, Alexander Depp, Olga Avrutina, Sven Dahms, Manuel Than, Martin Empting, Harald Kolmar, Engineering a constrained peptidic scaffold towards potent and selective furin inhibitors, *ChemBioChem* **2015**, *16*, 2441-2444.

Vanessa Siegmund, Stefan Schmelz, Stephan Dickgiesser, Jan Beck, Aileen Ebenig, David Fiebig, **Heiko Fittler**, Holm Frauendorf, Birgit Piater, Ulrich Betz, Olga Avrutina, Andrea Scrima, Hans-Lothar Fuchsbauer, Harald Kolmar, Locked by design: A conformationally constrained transglutaminase tag enables efficient site-specific conjugation, *Angewandte Chemie* **2015**, 45, 13420-13424.

Heiko Fittler, Olga Avrutina, Martin Empting, Harald Kolmar, Potent inhibitors of human matriptase-1 based on the scaffold of sunflower trypsin inhibitor, *J Pept Sci*. **2014**, *20*, 415-420.

Heiko Fittler, Olga Avrutina, Bernhard Glotzbach, Martin Empting, Harald Kolmar, Combinatorial tuning of peptidic drug candidates: High-affinity matriptase inhibitors through incremental structure-guided optimization, *Org Biomol Chem.* **2013**, *11*, 1848-1857.

Bernhard Glotzbach, Michael Reinwarth, Niklas Weber, Sebastian Fabritz, Michael Tomaszowski, **Heiko Fittler**, Andreas Christmann, Olga Avrutina, Harald Kolmar, Combinatorial optimization of cystine-knot peptides towards high-affinity inhibitors of human matriptase-1, PLoS One **2013**, 8, e76956.

Olga Avrutina, **Heiko Fittler**, Bernhard Glotzbach, Harald Kolmar, Martin Empting, Between two worlds: A comparative study on *in vitro* and *in silico* inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH, Org Biomol Chem. **2012**, 10, 7753-7762.

Tagungsbeiträge

12th German Peptide Symposium 2015 (Darmstadt)
Conjugation of high-affinity matriptase inhibitors onto the scaffold of an antibody Fc fragment
Heiko Fittler, Stephan Dickgiesser, Olga Avrutina, und Harald Kolmar

Bioorthogonal Chemistry - Meeting of the Biochemistry Division 2014 (Berlin)
Conjugation of high-affinity matriptase inhibitors onto the scaffold of an antibody Fc fragment
Heiko Fittler, Stephan Dickgiesser, Olga Avrutina, und Harald Kolmar

11th German Peptide Symposium 2013 (München)
High-affinity peptidic matriptase inhibitors through incremental structure-guided optimization
Heiko Fittler, Martin Empting, Olga Avrutina, Bernhard Glotzbach, und Harald Kolmar

Zusammenfassung

In dieser kumulativen Arbeit wurden durch strukturbasiertes Wirkstoffdesign neue Inhibitoren auf Basis des Gerüstes des *sunflower trypsin inhibitor-1* (SFTI-1) Peptids gegen krankheitsrelevante Serinproteasen generiert und untersucht.

Die erste Studie (Org. Biomol. Chem., **2012**, 10, 7753-7762) demonstriert den Nutzen von SFTI-1 als Gerüstmolekül für Modellierungen und Dockingversuche. Es wurde gezeigt, dass der Austausch eines wichtigen Strukturmotives von SFTI-1[*1*,*14*], seine Disulfidbrücke, durch unterschiedlich substituierte 1,2,3-Triazole nur dann gegen die Modellprotease Trypsin erfolgreich ist, wenn die neue Struktur mit der ursprünglichen nahezu identisch ist. Im Rahmen dieser Arbeit wurden die Ergebnisse einer *in silico* Studie mit den experimentellen Daten der Inhibitionsversuche verschiedener Peptide gegen die krankheitsbezogene Protease Matriptase-1 verglichen. Daraus ergab sich, dass beide Datensätze ein nahezu gleiches Ergebnis lieferten, nämlich eine geringe Affinität zu dem Zielprotein. Dies überrascht, da die Oberfläche von Matriptase-1 in der Nähe des aktiven Zentrums negativ und SFTI-1[*1*,*14*] positiv geladen ist. Um dies zu verstehen, wurden die berechneten freien Energien der Verbindungen aus den *in silico* und *in vitro* Daten miteinander verglichen, wobei sich diese als nahezu identisch herausstellten. Daraus lässt sich schließen, dass die geminderte Affinität durch negative entropische Beiträge der *C*-terminalen Region dafür verantwortlich ist. Folglich sollte eine *C*-terminale Verkürzung oder der Austausch von bestimmten Aminosäuren die Bindungsaffinität gegenüber Matriptase-1 erhöhen.

Aus den Daten der *in silico* Studie folgt, dass die monozyklische Variante von SFTI-1 (SFTI-1[*1*,*14*]) als Ausgangsmolekül für weitere Verbesserungen für potentere Matriptase-1-Inhibitoren eine gute Wahl ist. Aus der Kristallstruktur des Matriptase-1-SFTI-1-Komplexes ergaben sich drei Positionen für die Austausche, die keine Bindung mit der Protease eingehen. Daher wurden die Aminosäuren Ile7 und Ile10 gegen nicht-natürliche Aminosäuren mit Azidfunktionen in der Seitenkette ausgetauscht. Mit Hilfe der Kupfer(I)-katalysierten Azid-Alkin-Cycloaddition (CuAAC) wurde mit Hilfe unterschiedlicher Alkine eine Bibliothek von 22 Peptidmimetika synthetisiert. Diese Verbindungen mit 1,2,3-Triazolen wurden in einem Inhibitionstest gegen Matriptase-1 auf ihre Aktivität hin untersucht. Zusätzlich wurde Phe12 durch weniger raumfüllende kanonische Aminosäuren ersetzt. Nur die Austausche an den Positionen 10 und 12 führten zu einer Verbesserung der Inhibition gegen Matriptase-1 im Vergleich zu dem Wildtyp SFTI-1[*1*,*14*]. Die Triazolvariante mit einem Amin in der Seitenkette an Position 10 wurde durch die positiv geladenen kanonischen Aminosäuren Lysin und Arginin ersetzt. Erstaunlicherweise hatten beide Austausche eine bessere Inhibitionskonstante als ihr Vorgänger. Dabei war die Variante mit einem Arginin affiner als mit einem Lysin. Das Peptid, das beide Verbesserungen an den Positionen 10 und 12 in sich trägt, besitzt eine Inhibitionskonstante K_i von 11 nM (703 nM für den Wildtyp SFTI-1[*1*,*14*]) und wurde daraufhin als *SFTI-1 derived matriptase inhibitor-1* (SDMI-1) bezeichnet. Diese Variante beinhaltet nur kanonische Aminosäuren und ist somit leicht durch automatische Festphasensynthese zugänglich.

Die nächste Studie beschäftigte sich mit der Suche nach einer adressierbaren Position im Molekül für den Einbau von maßgeschneiderten funktionellen Gruppen. Diese wird z.B. für die Oligomerisierung des Inhibitors oder die Konjugation an ausgewählte Gerüstmoleküle wie einen Antikörper oder ein C4b-Bindeprotein (dort können alle sieben α–Helices adressiert werden) benötigt. Dafür wurden verschiedene Positionen auf ihre Austauschbarkeit untersucht, aber der jeder dieser Austausche führte zu einer erheblichen Verminderung der inhibitorischen Aktivität. Natürlich kann eine neue Gruppe einfach an den *N*-terminus gekoppelt werden. Die Verlängerung der Kette durch ein Fluorophor sorgte allerdings ebenfalls für einen starken Verlust der Inhibition (K_i = 328 nM). Andererseits hatte die Zyklisierung des *N*- und *C*-Terminus keine negative Auswirkung auf die Inhibition. Aus diesem Grund scheint die fehlende Ladung des freien Amins nicht das Problem zu sein, sondern die ungünstige Wechselwirkung des Fluorophors mit der Oberfläche der Protease. Deshalb wurde ein ε-Fmoc geschütztes Lysin an die erste Position eingebaut. Dadurch kann nach der Entschützung der Seitenkette der Einbau der gewünschten Funktionalität genutzt werden. Die Verbindung mit Lys1

besitzt eine Inhibitionskonstante K_i von 2,1 nM und ist damit potenter als ihre Vorgänger, sie wurde daraufhin als SDMI-3 bezeichnet. Die Kopplung von verschiedenen Reportermolekülen, wie Fluorophoren oder anderen Verbindungen, zeigte nur einen geringen Verlust an Affinität gegenüber Matriptase-1. Somit ließ sich SDMI-3 an verschiedene oligovalente Biomoleküle konjugieren, die zu Tetrameren im Falle des Fc-Fragmentes eines Antikörpers oder sogar zu einem Heptamer im Falle des C4b-Bindeprotein führte.

Die verbesserten Varianten aus den vorangegangen Studien wurden als Startpunkt für die Entwicklung neuartiger Furininhibitoren genutzt. Diese Protease steht im Zusammenhang mit Alzheimer, Krebs, Arteriosklerose und anderen Krankheiten. Der Wildtyp SFTI-1[*1*,*14*] zeigt nur eine moderate Inhibition von Furin (K_i = 35 μM), aber SDMI-3 erwies sich mit 24,1 nM bereits als potente Verbindung. Der Einbau der Furinsubstratsequenz (Arg-X-Arg/Lys-Arg↓) in SDMI-3 führte zu keiner Verbesserung der Inhibition. Interessanterweise erwies sich ein Arginin an der P1-Position in diesem Fall nicht optimal. Allerdings führte der Austausch gegen ein Lysin zu einer affineren Variante. Darüber hinaus sorgte der Austausch von neutralen gegen positiv geladene Aminosäuren für eine bessere Inhibition mit der negativ geladenen Oberfläche der Peptidase. Zusätzlich zeigte das *in silico* Modell keine ausgeprägte Wechselwirkung zwischen der *C*-terminalen Region und der Furin-Oberfläche. Daraufhin wurden verkürzte Varianten synthetisiert, bei denen die Aminosäuren 13-14 bzw. 12-14 fehlten. Beide Minimierungen hatten einen großen Einfluss auf die Aktivität der beiden Verbindungen mit Inhibitionskonstanten von 0,49 und 0,71 nM gegenüber Furin. Die aktivere Variante mit 12 Aminosäuren wurde daraufhin als *SFTI-derived furin inhibitor* (SDFI) bezeichnet.

Die neu, nach rationalen Kriterien designten Peptide SDMI-3 und SDFI sind vielversprechende Varianten, die in der Nuklearmedizin mit radioaktiven Markierungen zum Beispiel in der Tumordiagnostik eingesetzt werden könnten. Dafür wären die Verfahren der Positronen-Emissions-Tomographie (PET) und Einzelphotonen-Emissionscomputertomographie (SPECT) hervorragend geeignet.

Danksagung

Den folgenden Personen möchte ich meinen aufrichtigen Dank aussprechen, denn ohne sie wäre diese Arbeit nicht von Erfolg gekrönt.

Zu Beginn muss und will ich **Prof. Dr. Harald Kolmar** dafür danken, dass er mich in seinen Arbeitskreis aufgenommen hat und mir in meiner Zeit als Doktorand auch viele Freiheiten gelassen hat. Diese konnte ich dafür nutzen, um eigenen Projekten und Fragestellungen nachzugehen.

Desweiteren möchte ich bei meiner Korreferentin **Prof. Dr. Schmitz** und meinen beiden Fachprüfern **Prof. Dr. Buntkowsky** und **Prof. Dr. Biesalski** bedanken, dass Sie diese Aufgabe übernommen haben.

Olga Avrutina ist wahrscheinlich meine größte Förderin in diesem Arbeitskreis gewesen. Vielen Dank für deine Unterstützung und Anregungen bei meinen Projekten.

Meinen größten persönlichen Dank gilt **Martin Empting**, der mich damals in der Bachelorarbeit für die Wissenschaft begeistert hat. Ohne dich hätte ich wohl nie die Begeisterung und den Elan entwickelt, der für dieses Projekt von Nöten war.

Bei **Daichi Nasu** der mich die gesamte Zeit als Laborplatznachbar begleitet hat und mich immer während meiner Promotion unterstützt hat.

Bei **Stephan Dickgiesser** meinem Schreibnachbarn, für all das nützliche biologische Wissen das er mir vermittelt hat und dafür, dass er den zweifelsfrei besten Humor in dem Laden hatte (und das er immer für frische Luft gesorgt hat).

Bei **Stefan Zielonka** dem wahren Champ, den ich wohl nie wieder im Kicker besiegen werde. Ich hoffe wir werden noch einige Feste miteinander feiern dürfen. Brandon Boyd, wäre stolz auf dich.

Bei **Bernhard Valldorf** für die vielen Trainingseinheiten (leider meistens mit Handy) und für seine Kochkunst vor Lasertag, für die ich ihn sehr bewundere.

Desweiteren bei den beiden letzten überlebenden Chemikern **Chrischtl Uth** und **Sascha Knauer** für die angenehme Zeit im Labor.

Bei **Doreen Könning** für die vielen Scrubs-Zitate, die mich immer erheitert haben. Auch wenn ihr Faible für Faultiere etwas seltsam erscheint. Außerdem ist es bewunderswert, dass sie so schnell reden und gleichzeitig atmen kann.

Nicolas Rasche für die vielen Versuche die er für mich in seiner Zeit durchgeführt hat. Ein Mann der über Autos läuft - ein wahrer Supercop.

Natürlich bedanke ich mich auch bei der besten TA **Janine (Becker) Fritz**, die trotz einer fragwürdigen Namenswahl der Lichtblick eines jeden Morgens war.

Bei unserem unkündbaren Personal mit **Barbi Diestelmann**, **Andreas Christmann**, **Simone Bartl-Zimmermann** und **Cecilla Gorus**.

Bei **Vanessa Siegmund** die zwar immer im Raum war, wenn ein Gerät kaputt ging, aber trotzdem eine Bereicherung für diesen Arbeitskreis darstellt.

Bei **Julius Grezschik** der mit seiner zum größtenteils exzellenten Musik, die Stimmung immer wieder erheiterte. Hoch die Hände - Wochenende.

Natürlich bei den ganzen Leuten von Merck mit **Christian Schröter**, **Carolin Sellman**, **Simon Krah** und **Jan Beck**.

Außerdem möchte ich mich bei dem ganzen ehemaligen Chemikerteam mit **Michael Reinwarth**, **Sebastian Hörner** und **Sebastian Fabritz** bedanken.

Natürlich bei einigen tollen Personen, die mich in meiner Anfangszeit begleitet haben **Björn**, **Niklas**, **Thomas**, **Franzi**, **Alex** und **Bernhard**.

Genauso danke ich natürlich allen anderen aus dem **AK Kolmar**, die diese lange Zeit fröhlicher und angenehmer gemacht haben.

Zum Schluss gebührt mein Dank meiner Familie, die mich immer unterstützt und meinen Weg nie in Frage gestellt hat.

Table of content

1.....Introduction 1
1.1. Engineering of new drugs 1
1.2. Serine proteases of therapeutic relevance 3
1.2.1. Matriptase-1 5
1.2.2. Furin 7
1.3. Disulfide-stabilized scaffolds for drug design 8
1.4. Aims and scope 12
2.....Cumulative Part 13
2.1. Between two worlds: a comparative study on in vitro and in silico inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH 14
2.2. Combinatorial Tuning of Peptidic Drug Candidates: High-Affinity Matriptase Inhibitors through Incremental Structure-Guided Optimization 25
2.3. Potent inhibitors of human matriptase-1 based on the scaffold of SFTI 36
2.4. Engineering a Constrained Peptidic Scaffold towards Potent and Selective Furin Inhibitors 43
3.....Summary 48
4.....Literature 50
5.....Abbreviations 55
6.....Supporting Information 58
6.1. Supporting Information for chapter 2.1 59
6.2. Supporting information for chapter 2.2 69
6.3. Supporting Information for chapter 2.3 94
6.4. Supporting Information for chapter 2.4 101
7.....Appendix 122

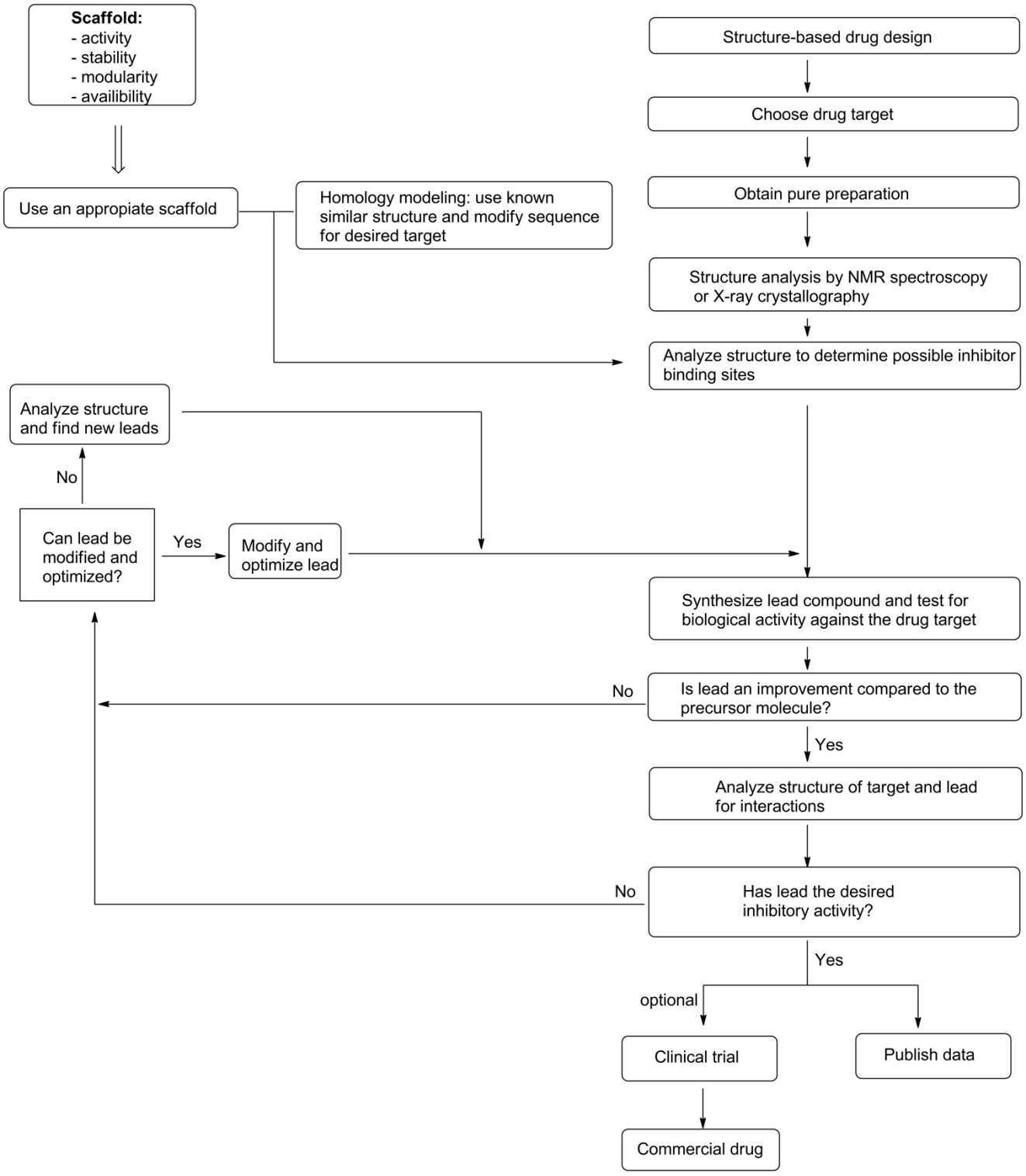
1. Introduction

1.1. Engineering of new drugs

Combating incurable diseases or inefficiently treated disorders requires intensive search and development of novel drugs. Different methods have been to date developed to access potent agents able to act against a challenging pharmacologic target. One of these strategies is the directed evolution based on multiple cycles of mutations in a gene or a protein in order to improve its kinetics, stability, or simply endowing it with a new function. Mimicking Nature that required millions of years to evolve a certain protein, directed evolution uses similar principles, thus tailoring an archetype according to the current needs within an incomparably shorter time. To this end, a template for the desired molecule, usually a protein, has to be chosen. Then, a library is generated using random substitutions, thus simulating an evolutionary selection. Different techniques are available to introduce these random mutations, among them the error-prone PCR of the DNA,[1-2] randomized oligonucleotides,[3-5] somatic hypermutation[6], deficient polymerases,[7-8] or DNA shuffling.[9] Obviously, no structural information on the target is required in this method, and advanced technologies have been developed for library screening, e.g. the bacterial,[10-11] phage,[12-13] mRNA,[14-15] ribosome,[16-17] or yeast display.[18-19] However, the huge size of the generated libraries often causes problems with handling and efficient screening within a tolerable time.

Rational design, being a more sophisticated, regulated, and target-oriented approach is a viable alternative to directed evolution. However, the structure of a target is a requirement of the method rather than its prerequisite. Based on structural information obtained from reliable analytic methods, a process called the structure-based drug design can be applied that makes use of *in silico* experiments. The ultimate goal of the method is to obtain binders against a target protein and tailor them towards desired potency. Therefore, as a first step the target has to be identified, usually upon linkage with a particular disease-related pathway. Afterwards, the target must be produced in an amount required (ideally) for the characterization with NMR spectroscopy or X-ray crystallography, and for activity assays. Structural information allows for modeling and docking experiments in order to reveal an appropriate scaffold. It should possess a moderate activity towards the target, combined with synthetic availability, stability, and certain modularity. Through the modeling process, substitutions of amino acids can be done, and binding affinity can be calculated from docking experiments. The next step is the verification of the *in silico* data. Therefore, the new lead variants have to be synthesized and tested for activity towards their drug target. If no enhancement of activity compared to the parent structure is observed, the iterative process of optimizing the lead molecule has to be continued, either by launching further modeling attempts or switching to a better design scaffold.[20]

If an enhancement of the affinity towards the drug target is observed, it has to be compared with the desired one. In the case the binding is not sufficient, the cycle of the lead optimization is repeated until the desired inhibitory activity has been obtained. The final compound can optionally be tested in a clinical trial and, brought to the market (Scheme 1).



Scheme 1: Algorithm of structure-based design of protease inhibitors (scheme modified from Anderson).[20]

1.2. Serine proteases of therapeutic relevance

In Nature, endopeptidases catalyze proteolytic cleavage of amide bonds within certain peptidic substrates. The Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) has defined proteases as "hydrolases acting on peptide bonds".[21] Approximately 500 genes in the human body encode for various types of proteases. Interestingly, peptidases make up about 2 % of proteins in all organisms,[22] with more than 4000 representatives being listed in the MEROPS database.[23] Each of these enzymes possesses different amino acids in the catalytic center, which promote the hydrolytic cleavage of a peptide bond in an exergonic reaction. This process, though being thermodynamically favored, has been never observed in pure water ($\tau_{1/2}$ = 10-1000 years) and a catalyst is required to accelerate it. This could be explained taking into consideration that a carboxamide group possesses partial double-bond character due to delocalization of the nitrogen lone pair into the carbonyl group. In the presence of a suitable biocatalyst degradation of a peptide bond reaches completion within milliseconds.

To date, seven families of proteolytic enzymes are listed in the MEROPS database (Table 1). However, for some proteinases the crystal structure, *ergo*, the catalytic mechanism remains unknown, and several members can be attributed to multiple families.

Table 1: Overview of the different families of peptidases, with their corresponding nucleophile and amino acids in the catalytic center.[24-29]

Type of protease*	Nucleophile	AA in the catalytic center	Presence in the human body
Asparagine (N)	water	Asp, Asn	0
Aspartic (A)	water	Asp, Asp	21
Cysteine (C)	thiol	His, Asp, Cys	149
Glutamic (G)	water	Glu, Gln	0
Metallo (M)	water	Glu, His	194
Serine (S)	hydroxyl	His, Asp, Ser	176
Threonine (T)	hydroxyl	Thr, Thr	28

*The alphabetic acronym for the type of a protease is not linked to a one-letter amino acid code.

Diverse factors influence the activation of peptidases, among them specific regulators, inhibitors, and pH of the reaction milieu. Being unregulated, endopeptidases, e.g. the matrix metalloproteinases (MMPs), can trigger certain diseases, among them arthritis, inflammation, or cancer.[26] Interestingly, the MMPs are the only proteases requiring a metal ion, usually zinc, in the active site in order to activate water as a nucleophile for the cleavage of the peptide bond. Some other enzymes are overexpressed in course of pathologic processes resulting in uncontrolled and undesired cleavage activity as well as activation of unwanted pathways. Therefore, control over the activity of endopeptidases plays an important role in modern healthcare.

Trypsin, a pancreatic enzyme found in many vertebrates, is the most studied serine protease. Known to man since millennia, this protease is responsible for digestion, blood coagulation, fibrinolysis, development, fertilization, apoptosis, and immunity.[30] In the following, the mechanism of the catalytic triad in serine proteases (which amount to one-third of all endopeptidases) will be explained in detail using trypsin as a typical representative.

Being similar for almost all proteases, the catalytic mechanism starts with a nucleophilic attack on the carbonyl carbon of an amide bond.[31] In serine proteases, the required nucleophile is provided by the side chain of Ser^{195} in the active site, with the proton of the hydroxy group linked to the imidazole ring of His^{57} *via* a hydrogen bond (Figure 1).

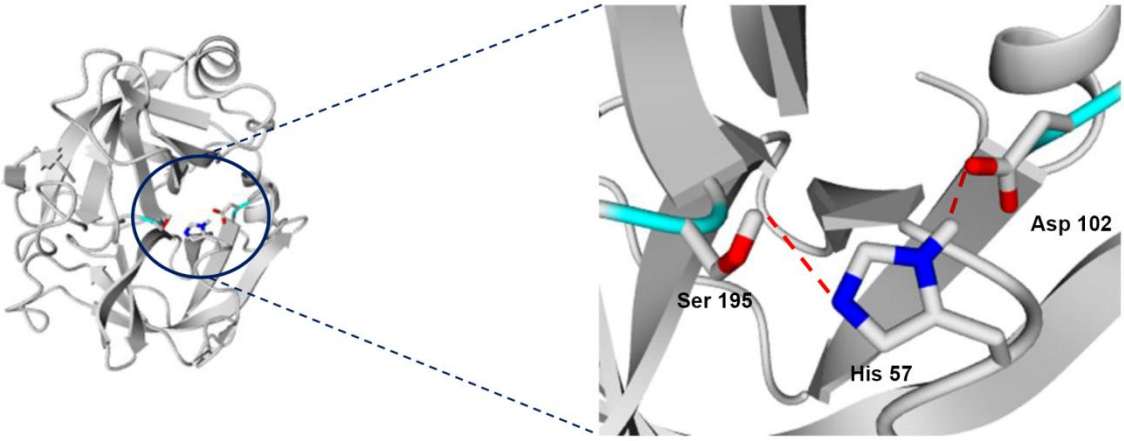


Figure 1: Secondary structure of human matriptase-1 (PDB: 3P8F) (left) with a close-up of the three amino acids at the active site (right).

The NH group of the imidazole ring is also linked to a carboxylic group of Asp[102]. These three amino acids, referred to as the catalytic triad, compose a highly coordinated structure in the very heart of an enzyme and are the key players in the hydrolytic cleavage of peptide bonds. The nucleophilic oxygen of the hydroxy group of Ser[195] attacks the carbonyl carbon of the scissile peptide bond (Figure 3a). This process is catalyzed by the imidazole group from His[57] serving as general base, leading to a tetrahedral intermediate stabilized by NH-groups on the protein's surface (Figure 3b), the so-called oxyanion hole (Figure 2). A proton from the positively charged imidazole ring is transferred to the primary amine of the scissile peptide bond, and the intermediate degenerates to an acyl-enzyme and an amine product. The acyl-enzyme gets deacylated upon a nucleophilic attack by a water molecule, which gets deprotonated first by the nitrogen of the imidazole group (Figure 3c). This second addition-elimination reaction leads to a regeneration of the catalytic triad and the process can continue with a new peptide bond to cleave (Figure 3d-e).

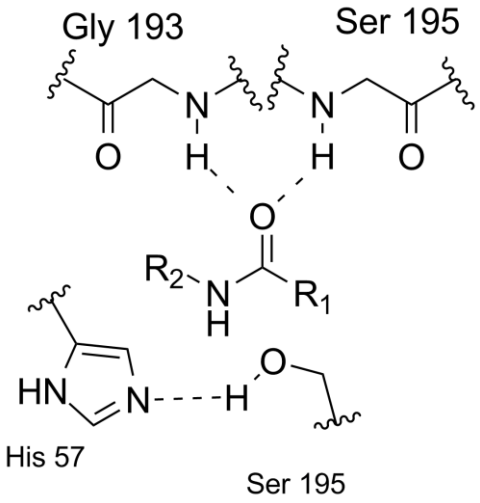


Figure 2: Depiction of the oxyanion hole of chymotrypsin. The occurring tetrahedral intermediate will be stabilized by the amine groups of Gly[193] and Ser[195].

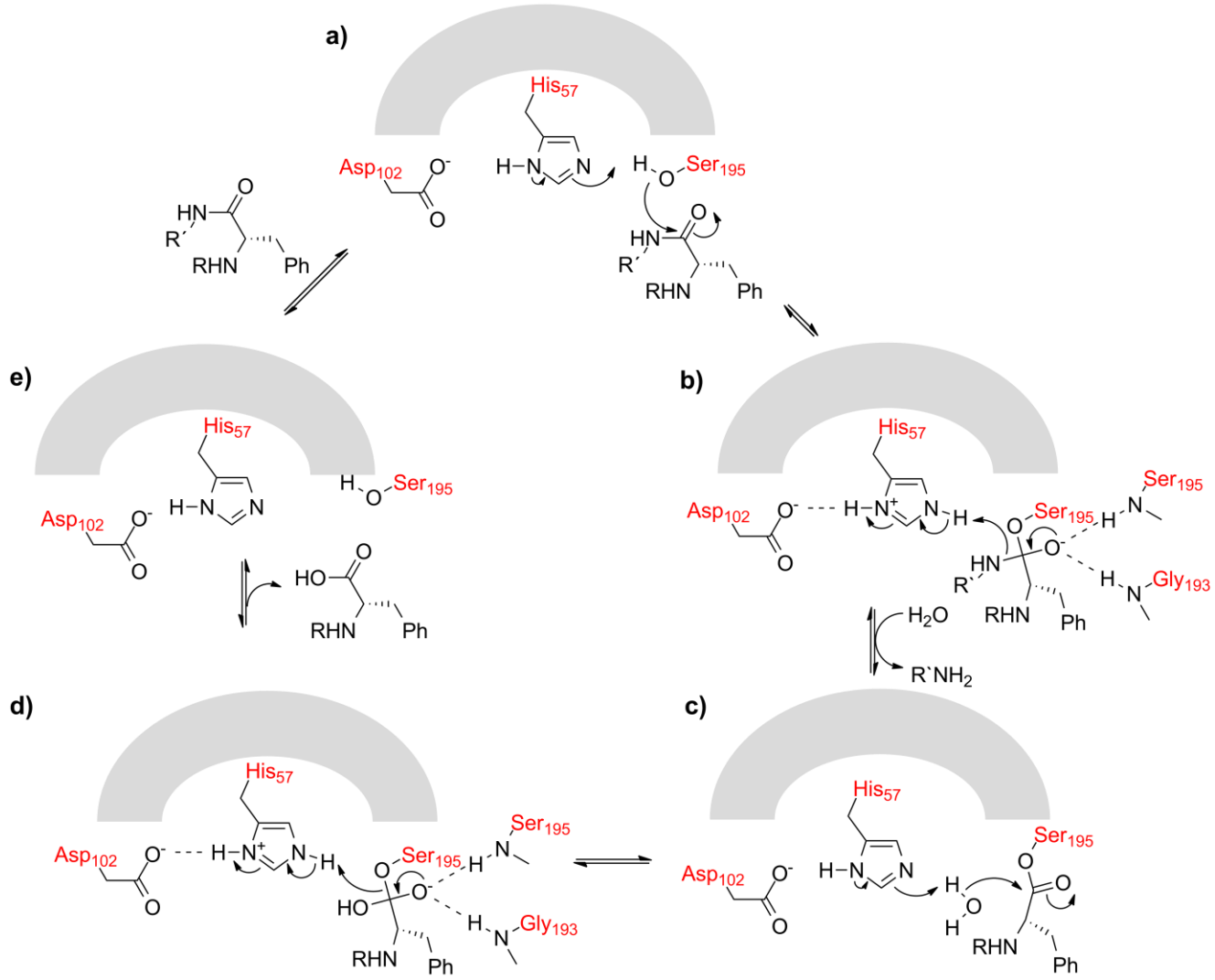


Figure 3: Catalytic mechanism of a serine protease. The carbonyl carbon of the substrate is attacked by the hydroxyl of serine, followed by the cleavage of the amine. A water molecule adds to the carbonyl carbon and the cycle can be repeated.

1.2.1. Matriptase-1

In proteases, the mechanism of the catalytic triad provides a major information about the cleavage of a particular substrate. Nevertheless, for the design of an efficient protease inhibitor, more precise information on the active site's surroundings is required, e.g. amino acid composition, surface charge, and architecture of the binding pockets at sufficient resolution. Therefore, for a structure-based drug design structural insights into a target of interest are of high priority.

Matriptase-1 is a member of the family of type-II transmembrane serine protease (TTSPs) and is also referred to as MT-SP1, epithin, serine endopeptidase SNC19, serine protease TADG-15, and CAP3.[23] Discovered in 1993 in human breast cancer cells, it is still in the focus of numerous disease-related studies.[32] Indeed, matriptase-1 is associated with breast[33-34] and ovarian cancer,[35] atherosclerosis,[36] osteoarthritis,[37] chronic lymphocytic leukemia,[38] and other tumors.[39-42]

Structurally, matriptase-1 comprises the extracellular stem region including a single sea urchin sperm protein (SEA), two urchin embryonic growth factor and bone morphogenic protein-1 domains (CUB). Its catalytic triad is located *C*-terminally; four putative N-linked glycosylation sites are disposed at positions 109, 302, 485, and 772. The 80-90 kDa glycoprotein is anchored onto the cellular membrane *via* its *N*-terminus, expressed in several epithelial tissues, and has pleiotropic effects on development, cell-cell adhesion and homeostasis (Figure 4).[43-44]

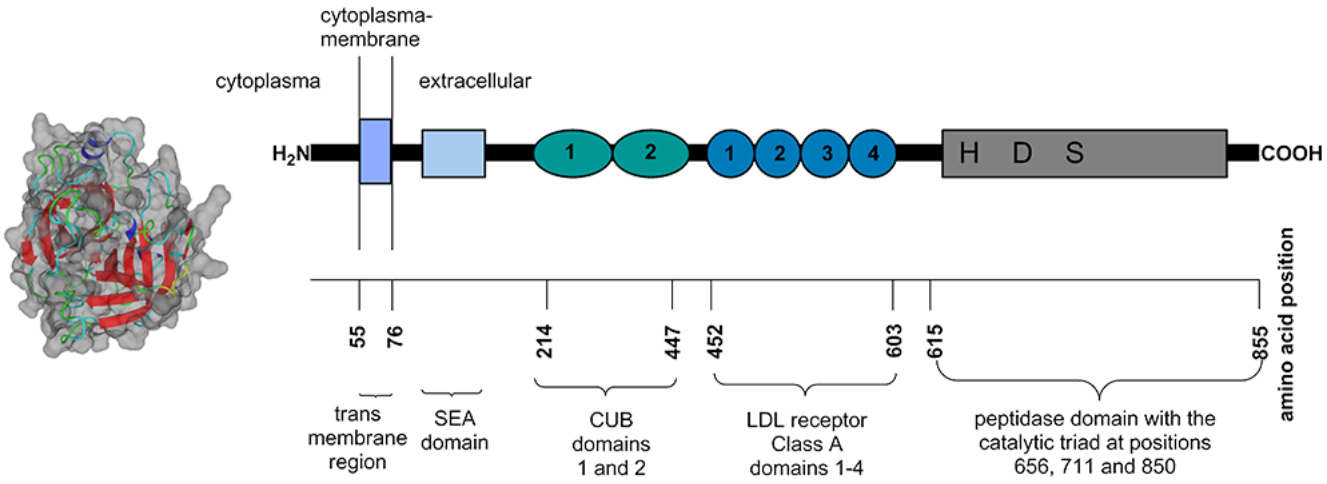


Figure 4: Left: Crystal structure of matriptase with elements of secondary structure indicated (PDB: 3P8F). Right: Multidomain structure of matriptase-1 comprising the *N*-terminal cytoplasmic domain and membrane, the extracellular domain containing the stem region (including SEA and two CUB domains) and four LDLRA (low-density lipoprotein receptor domain) domains, as well as the catalytic triad at the *C*-terminus (figure modified from K. Uhland).[45]

Matriptase-1 is naturally controlled by its Kunitz-type serine protease inhibitor hepatocyte growth factor activator inhibitor-1 (HAI-1) which plays a crucial role in the activation of this protease after its expression as a zymogen. Without HAI-1 the activity of MT-SP1 drops and the overall surface expression is reduced.[46-47] The crystal structure of matriptase-1 in complex with HAI-1 (PDB: 4ISO) gives insights into the binding geometry, with the Kunitz domain 1 (KD1) enabling the interaction. This motif has a Gly-Arg-Cys-Arg↓-Gly binding sequence between the P4 to P1' residues (nomenclature of Schechter and Berger[48]).[49] Interestingly, the substrate sequence required for the recognition by matriptase-1 has been defined as Arg/Lys-X-Ser-Arg↓-Ala (X stands for any non-basic amino acid).[50]

Being able to cleave and activate numerous substrates and enzymes, matriptase-1 represents a challenging target for drug development. Thus, *in vitro* experiments verified that it is recruited in activation of the zymogen of the hepatocyte growth factor (pro-HGF) and the urokinase-type plasminogen activator (pro-uPA).[50-51] Both enzymes are involved in the process of invasive growth leading to cell proliferation, which is required for organ regeneration and maintenance but also in tumor growth and metastasis.[45]

Over the past decade numerous studies aimed at the development of synthetic matriptase-1 inhibitors have been conducted. All these small molecule compounds shared the positively charged residues around the P1 position. Using bis-benzamidines as a scaffold for the structure-based drug design, potent inhibitors were developed combining a K_i of 208 nM with good selectivity profile against the urokinase plasminogen activator (uPA) and thrombin.[52] Steinmetzer *et al.* tailored a library of inhibitors based on bis-basic secondary amides of sulfonylated 3-amidinophenylalanine for better interaction with matriptase-1 subsites, resulting in a compound with a K_i of 3.9 nM.[53] This molecule was further improved by replacement of the *C*-terminal basic group with a more neutral one, resulting in a K_i of 0.43 nM.[54] The most potent compound, to our knowledge, is an irreversible peptidic inhibitor with a K_i of 11 pM which possesses a ketobenzothiazole group at the *C*-terminus, thus making matriptase-1 no longer able to cleave the peptide (Figure 5).[55] Although a number of very potent inhibitors of MT-SP1 has been developed over the last years,[56-61] there is still room for improvement regarding a prolonged half-life in the human body.

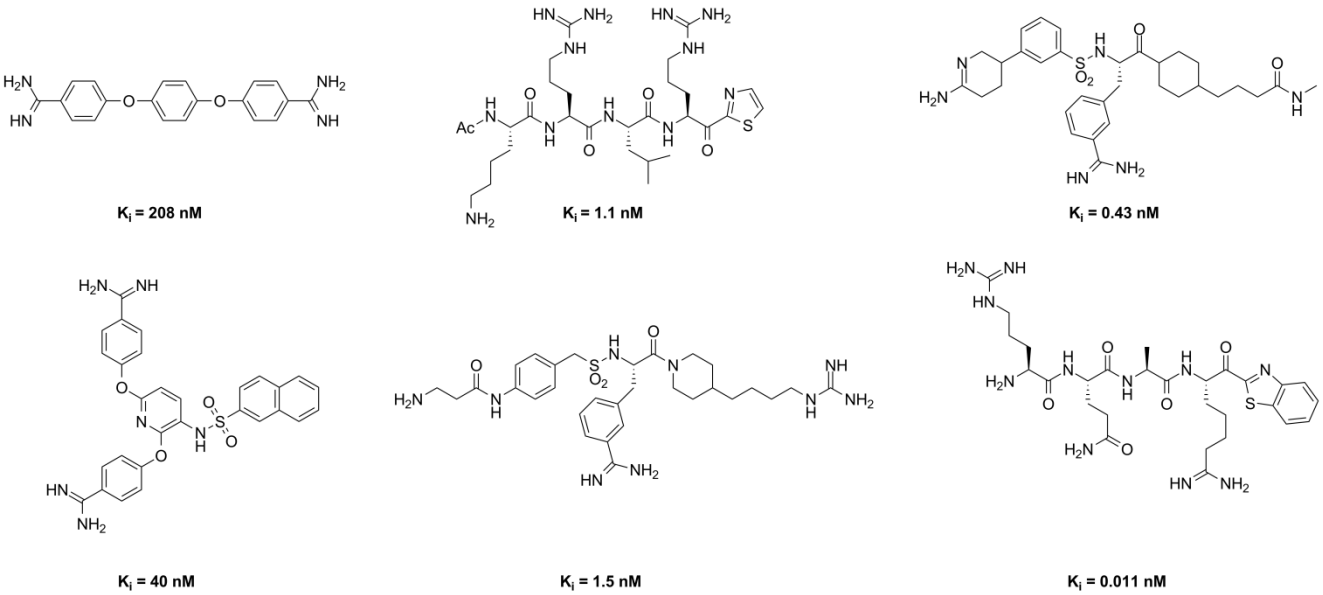


Figure 5: Selected potent matriptase-1 inhibitors.[52-55, 59, 62]

1.2.2. Furin

The type-I transmembrane serine protease furin is a member of the proprotein convertases (PCs) family, together with six other members (PACE4, PC1/3, PC2, PC5/6, PC7, and PC4). This subtilisin-like protease is of growing interest due to its involvement in cancer, Alzheimer's disease, viral infection, osteoarthritis, atherosclerosis, and long-term pathogen infections.[63-66] Discovered in 1986, the strictly calcium dependent endoprotease is expressed in all cells and mainly active in the trans-Golgi network. It contains a signal peptide that directs the expressed protein to the secretory pathway. The prodomain functions as a chaperone for proper folding and as an intramolecular inhibitor in order to control the activity of furin. It gets cleaved at a neutral pH and a sufficient calcium concentration. Other than in matriptase-1, the catalytic domain is located *N*-terminally. The P-domain is responsible for the calcium and pH dependence and enzyme stability. There are three possible glycosylation sites located in the catalytic and P-domains. Finally, the transmembrane domain anchors the endopeptide to the plasmatic membrane (Figure 6).[64, 66-67] In 2003 Henrich *et al.* solved the crystal structure of mouse furin inhibited by the decanoyl-Arg-Val-Lys-Arg-chloromethylketone (dec-RVKR-cmk) (PDB: 1P8J).[68]

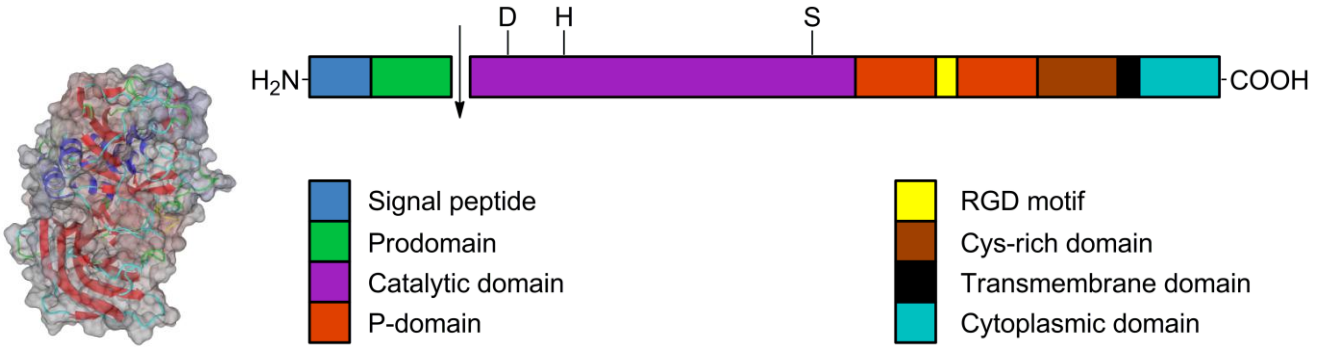


Figure 6: Left: Crystal structure of furin with elements of secondary structure indicated (PDB: 1P8J). Right: Sketch of multidomain structure of furin with its catalytic triad shown in one-letter amino acid code.[64, 66-67]

Shortly after the discovery of furin, various kinds of its inhibitors were synthesized and their activity was evaluated. The first generation of inhibitors were the irreversible decanoyl-peptidyl-chloromethyl

ketones with single-digit nanomolar K_i's.[69-70] One of the first protein-based inhibitors was the α_1-antitrypsin Portland (α_1-PDX), modified from the naturally occurring α_1-antitrypsin serum protease inhibitor with a K_i of 600 pM, by simply inserting the furin substrate sequence RIPR in the serpin reactive site.[67, 71] Screening for oligopeptide binders against furin revealed polyarginine peptides, with nona-L-arginine being the most potent one (K_i of 42 nM).[72] As the surface of furin is negatively charged, positively charged residues in the binding loop of a potential inhibitor are required. Kacprzak *et al*. were able to improve the inhibitor affinity by using D- instead of L-polyarginines, resulting in nona-D-arginine with a K_i of 1.3 nM.[73] The development of a 2,5-dideoxystreptamine library demonstrated great potential of this scaffold (the best binder possessed a K_i of 6 nM).[74] A different small-molecule library of inhibitors was generated based on 4-amidinobenzylamide with a K_i of the best compound of 0.81 nM.[75] In a following study, the compounds were further improved resulting in an affinity to furin of 8 pM (Figure 7).[76]

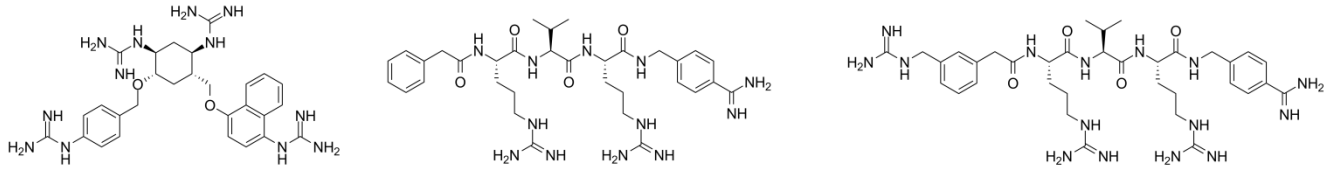


Figure 7: Selected potent furin inhibitors.[74-76]

1.3. Disulfide-stabilized scaffolds for drug design

During the evolution, Nature developed a vast number of scaffolds – molecular platforms serving as starting points in structural design towards generation of new functions and properties - which differ in size, architecture and complexity (Figure 8). Although these frameworks often diverge considerably from each other, they possess several common features: stability, activity, availability, and modularity. Among these scaffolds, disulfide-stabilized peptidic architectures can be indicated as a separate group, despite they differ in morphology, biologic function, and cystine fold as a prerequisite for constitution and stability.

The biggest disulfide-stabilized scaffolds applied in immunotherapy are the antibodies or immunoglobulins (Igs). They are divided into five groups and consist of IgG, IgA, IgM, IgD and IgE, distinguished by their size and structure. Thus, a typical IgG molecule is a tetramer consisting of two identical light and heavy chains comprising repeated structural motifs (homology regions).[77] They bear two antigen binding sites (Fabs) and one constant fragment (Fc). Their complementarity determining regions (CDRs) are responsible for the diversity of the antibodies and for binding the epitope of a particular target. The disulfide bridges provide structural integrity and their number differs from 2 for IgD to 12 for IgG3. Antibody-based cancer therapy is to date the main driver of sales at the pharmaceutical market.[78]

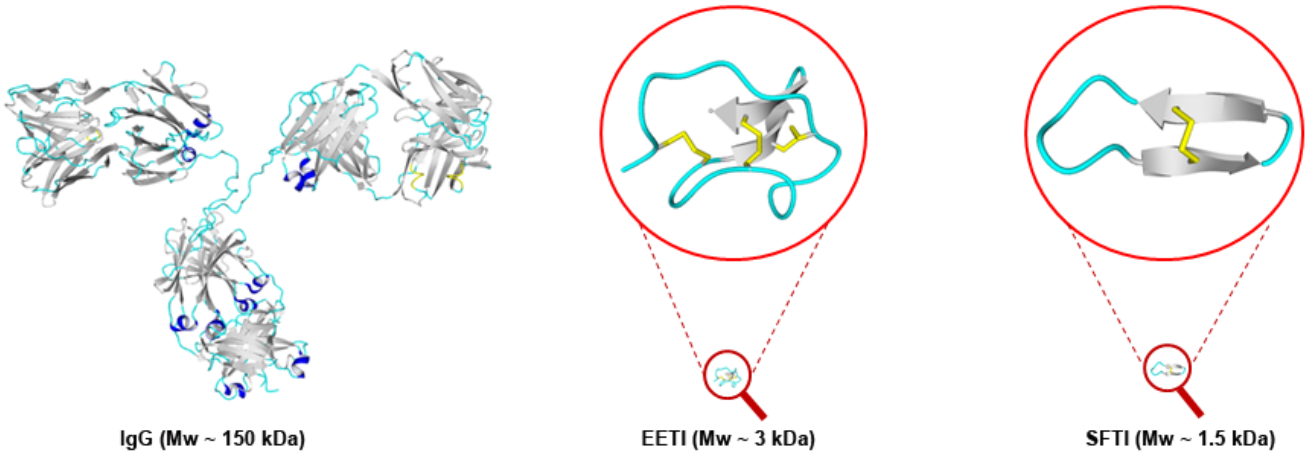


Figure 8: Comparison of the molecular weight of an antibody (PDB: 1IGT), EETI (PDB: 2IT7), and SFTI-1 (PDB: 1SFI).

Another group of disulfide-stabilized molecular scaffolds have been developed using the representatives of a large family of the so-called cystine-knot peptides. These unique architectures possess nearly all the properties of proteins, except for the molecular size. Indeed, they have a number of beneficial features for drug design, such as temperature and protease stability, synthetic accessibility, and modularity.[79] These molecules, also known as *knottins,* consist typically of 30-50 amino acids and have a core of antiparallel β-strands which is stabilized by at least three disulfide bonds (Figure 8, Figure 9a).[80] To date, a broad spectrum of knottin species has been identified, which possess many-sided biological activity, e.g. insecticidal, inhibitory, anti-HIV, cytotoxic, antimicrobial, or hormone-like.[81] In these molecules, the cysteines are always connected in the same way, the first with the fourth, and the second with the fifth one. The disulfide bond between the third and the sixth cysteine penetrates the ring made by other two cystines, thus forming a geometrical pseudo-knot. Sometimes the cystine-knot structure is additionally stabilized by head-to-tail macrocyclization; these peptides are called cyclotides. Although it is still questionable how this macrolactamization takes course in Nature, it is believed to provide enhanced bioactivities by increasing stability and lowering flexibility; better resistance against proteolytic cleavage is another advantage.[82] In Figure 9, the structures of cyclic and open-chain knotted miniproteins are depicted.[83]

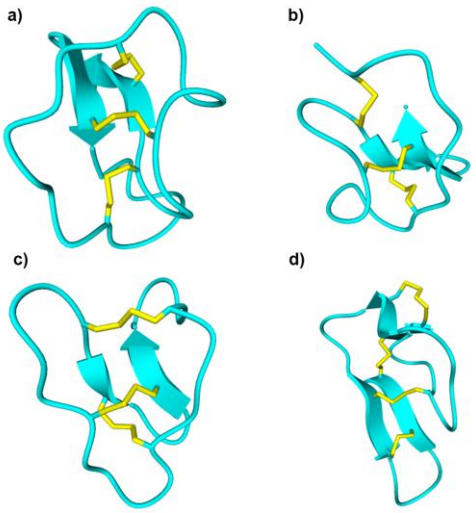


Figure 9: Depiction of different disulfide-rich miniproteins. a) MCoTI-II (PDB: 1HA9), b) EETI (PDB: 2IT7) c), Kalata B1 (PDB: 2LUR) d) AgRP (PDB: 1MR0).[80]

Several representatives of the knottin family have got into research focus due to their potential as drugs and diagnostic agents,[84] among them miniproteins from the fruit of *Momordica cochinchinensis* (trypsin inhibitors [MCoTI's]),[85] the seeds of the squirting cucumber *Ecballium elaterium* (trypsin inhibitor [EETI]),[82] an insecticide from the plant *Oldenlandia affinis* (kalata B1),[86] a fragment of the Agouti-related protein (AgRP), numerous toxins from animal species, and many others (Figure 9).[80, 87] They are able to keep their functional and structural integrity even upon exposing to harsh conditions like strong acids and bases, high temperature and proteolytic enzymes.[79-80, 88] Their outstanding stability carries hopes for potential oral application, which often is a drawback for other protein-based therapeutics. In 2007 Greenwood *et al.* demonstrated the capability of the miniprotein MCoTI-II to internalize into a breast cancer cell line and macrophages.[89] The knottins were used in grafting experiments with bioactive peptides inserted in the loops of the cyclotides. Being still able to

fold towards a bioactive conformation; the resulting molecules demonstrated enhanced affinity towards their targets.[90-91] Glotzbach *et al*. showed the viability of this grafting method by generating yeast surface-displayed libraries of MCoTI with a diversity of 2 x 10^7, with the best candidates possessing inhibitory activity against therapeutically relevant protease matriptase-1 in a subnanomolar range.[56] As the design of the library can be combined with a rational approach taking in consideration that some amino acids are more favorable at certain positions, the probability of getting improved binders is increased.[92-93]

The members of Bowman-Birk inhibitor (BBI) family possess intrinsic antiproteolytic activity. In plants the BBIs are involved in the defense mechanism, thus bearing potential as cancer chemopreventive.[94] They all share the same rigid and conserved canonical loop comprising an extensive hydrogen-bond network, a disulfide bridge, and a *cis*-proline motif (Figure 10).[95] Some inhibitors benefit from their flexibility as a key attribute; in the case of the BBIs the rigidity helps to lock the inhibitor loop in a complementary conformation to the enzyme.[96]

The crucial role of the mentioned elements has been proven in a number of experiments. Thus, the requirement of the *cis*-Pro at the P3' position (nomenclature of Schechter and Berger[48]) was confirmed as a dramatic loss of binding affinity was observed upon a substitution of this residue.[94-95, 97] Interestingly, the P1 residue important for binding at the active site in the S1 pocket of the protein plays no role in the rigidity of the inhibitor construct,[96] concluding the possibility to change the specificity of a BBI towards an alternative target protease without loss of its characteristic conformational rigidity. Most of the members of the family have a molecular weight in the range of 6-8 kDa and are stabilized by 7 disulfide bonds. Many of them are capable to bind two different proteases at the same time, usually trypsin and chymotrypsin.[98]

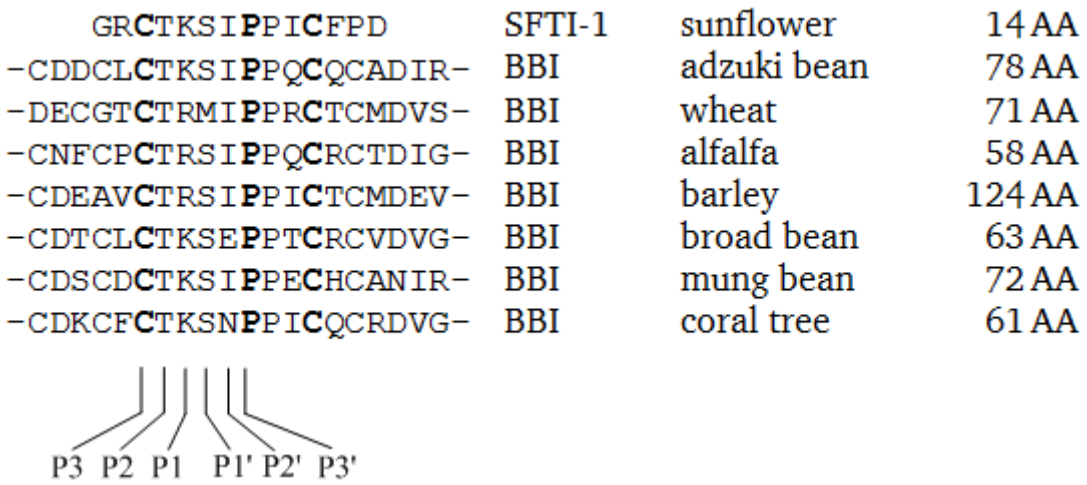


Figure 10: Active site sequence of Bowman-Birk inhibitors in comparison. A conserved disulfide bond formed by two cysteines (bold), and a *cis*-Pro at the P3' position (bold) are common in all BBIs. Origin and length of the molecules are also stated (figure modified from Daly *et al.*).[94]

One of the smallest and most potent BBIs known is the sunflower trypsin inhibitor-1 (SFTI-1), which was isolated 1999 by Luckett *et al*. from seeds of *Helianthus annus*.[99] To date, a detailed analysis of its structure and inhibitory activity towards different proteases has been done.[98, 100-104] In contrast to other inhibitors of the BBI family, comprising typically 60-70 residues, SFTI-1 contains only the rigid canonical loop of 14 amino acids (Figure 10).[99] The backbone of SFTI-1 is cyclic and additionally stabilized by a disulfide bridge which is responsible for the general rigidity of the peptide (Figure 11).[105-106]

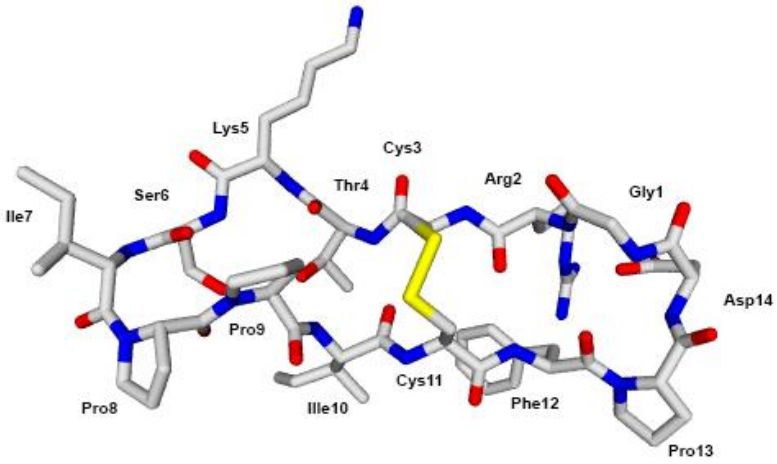


Figure 11: Depiction of the sunflower trypsin inhibitor-1 (SFTI-1) (PDB: 1SFI). Blue: nitrogen, red: oxygen, yellow: sulfur, grey: carbon, hydrogen atoms are omitted for clarity.

As the name implies, SFTI-1 is a very potent trypsin inhibitor with a K_i of 0.1 nM. During the last decade, numerous attempts have been made to improve or tailor the affinity and selectivity of SFTI towards other targets. To this end, it is important to identify, on the one hand, the residues which are involved in the interaction with the protein and, on the other hand, those which are essential to keep the conformation of the backbone intact. Therefore, a mutation of every amino acid against an alanine (an Ala-scan) became an appropriate method which allowed for the identification of the key amino acids within the SFTI-1 scaffold.[94, 107-108] Thus, three amino acids, Gly1, Pro9, and Pro13, were only marginally involved in the interaction with trypsin, while Lys5 and Pro8 were found indispensable. Indeed, the Lys5 is required for the optimal fit in the S1 pocket of trypsin, and *cis*-conformation of Pro8 is a fundamental demand for a BBI fold. Other elements of the sequence play their role upon binding, being however a second-line participants in the inhibitory process.

The role of the disulfide bond as well as its substitution against a more redox stable connectivity has been studied as well.[109] Thus, Zablotna *et al*. investigated the effect of the head-to-tail cyclization and the cystine motif and found the bicyclic SFTI-1 being the most potent trypsin inhibitor followed by the acyclic disulfide-bridged variant. Not surprisingly, the inhibitory activity of the peptide lacking the cysteines as a result of their substitution by α-aminobutyric acid decreased most of all.[110] In a different approach, the sulfur was substituted against the isosteric selenium by incorporation of selenocysteine, and the K_i was only 3-fold decreased with the new diselenide bridge.[111] Liu *et al*. used a more sophisticated approach and exchanged the disulfide against a methylenedithioether bridge as a more redox stable linker. The inhibition of the serine protease matriptase was examined, and the engineered peptide showed nearly the same K_i as the parent SFTI-1.[112] Finally, Empting *et al*. showed the utility of 1,5-disubstituted 1,2,3-triazoles as viable disulfide surrogates.[109]

Additionally, the introduction of non-natural amino acids or other non-canonic compounds at the desired positions has been studied. As the repertoire of synthetic building blocks bearing different side chains is nearly unlimited, the incorporation of these elements seems intriguing as it can not only lead to novel molecules with enhanced potency, but provides insights in the structure-activity relationship (SAR) for the new variants.[100, 112] An alternative approach makes use of the application of N-substituted glycine residues, the so-called peptoids. As these amino acid analogs are fully protease-resistant,[113] their application could significantly increase the stability of a potential drug. However, the Lys5 interacts nearly perfect with the S1 pocket of the enzyme suggesting that substitution at this position could lead to decrease in inhibitory activity.

For diagnostic purposes, e.g. imaging, the derivatization of peptide with markers like fluorophores or radiolabels is of high interest. Several groups demonstrated that the labeled peptide still had a very potent inhibition of the target protease.[114-116]

1.4. Aims and scope

During the last decades, natural and engineered compounds possessing peptidic or peptide-like structure emerged as a challenging and rapidly commercializing field. In 2011, the net sales of peptides worldwide were estimated to 14.7 billion USD, [117] clearly demonstrating the importance of these molecules for the drug market today. Indeed, compared to small molecules and full-size proteins, peptidic drugs have the advantage of being highly selective, efficacious, relatively safe, and well tolerated.[118]

It is estimated that up to the year 2030 the deaths caused by cancer will be around 11.5 million people. To date, monoclonal antibodies which are able to distinguish and eliminate tumor cells with high selectivity are often the therapy of choice. However, their high molecular weight, low tissue penetration, as well as poor cellular uptake significantly limit their potential. In contrast to antibodies, intrinsic solubility, favorable pharmacokinetics, and distinct tissue distribution reside in peptides, which enables both efficient uptake and fast elimination.[119]

Several anticancer peptides have been to date either isolated from animal sources or generated *via* library screening, e.g. by phage display.[120-123] *De novo* drug design is another approach to develop peptides with tailored, e.g. anti-tumor, properties. To this end, bioinformatics and computer-aided design (CAD) are used, with precise knowledge on both target and lead being an indispensable prerequisite.[124-125] Generally, the structure-based design is a powerful tool towards the development of novel and optimization of known inhibitors of certain therapeutically-relevant proteases. With the information on parent scaffold and the protease of interest, binding properties of an inhibitor can be improved or tailored according to the instant needs by means of molecular engineering.

Sunflower trypsin inhibitor-1 (SFTI-1) is a disulfide-stabilized cyclic peptide which emerged as a viable tool for the generation of novel protease inhibitors.[61, 126-135] Engineering this scaffold towards potent inhibitors for different protease targets of therapeutic relevance is the main goal of this thesis. Therefore, the objectives of this work can be summarized as followed:

- Comparative study of *in vitro* and *in silico* inhibition of trypsin and matriptase-1 by SFTI-1[*1*,*14*] derivatives bearing disulfide-mimicking disubstituted 1,2,3-triazoles;
- Synthesis of SFTI-1[*1*,*14*]-triazolyl derivatives with improved affinity to therapeutically relevant protease matriptase-1, by amino acids replacement within negligibly interacting regions;
- Analysis of contributions from each beneficial substitution and their combination towards improved activity;
- Tailoring of inhibitors towards further modifications upon installation of an additional addressable site for orthogonal transformations;
- Oligomerization of the matriptase-1 inhibitors using certain biomolecular scaffolds;
- Structure-based design of furin inhibitors based on the scaffold of SFTI-1[*1*,*14*].

2. Cumulative Part

This section contains the following articles that have been published in peer-reviewed journals.

2.1 Olga Avrutina, **Heiko Fittler**, Bernhard Glotzbach, Harald Kolmar, Martin Empting, Between two worlds: a comparative study on in vitro and in silico inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH. *Org. Biomol. Chem*. **2012**, 10, 7753-7762.
- Reproduced by permission of the Royal Society of Chemistry (RSC).

2.2 **Heiko Fittler**, Olga Avrutina, Bernhard Glotzbach, Martin Empting, Harald Kolmar, Combinatorial Tuning of Peptidic Drug Candidates: High-Affinity Matriptase Inhibitors through Incremental Structure-Guided Optimization. *Org. Biomol. Chem*. **2013**, 11, 1848-1857.
- Reproduced by permission of the Royal Society of Chemistry (RSC).

2.3 **Heiko Fittler**, Olga Avrutina, Martin Empting, Harald Kolmar, Potent inhibitors of human matriptase-1 based on the scaffold of sunflower trypsin inhibitor. *J Pept Sci*. **2014**, 6, 415-420.
- Reproduced by permission of John Wiley and Sons.

2.4 **Heiko Fittler**, Alexander Depp, Olga Avrutina, Sven O. Dahms, Manuel Than, Martin Empting, Harald Kolmar, Engineering a Constrained Peptidic Scaffold towards Potent and Selective Furin Inhibitors. *Chembiochem* **2015**, *16*, 2441-2444.
- Reproduced by permission of John Wiley and Sons.

Additionally, the following articles have also been published between 2012-2015.

Bernhard Glotzbach, Michael Reinwarth, Niklas Weber, Sebastian Fabritz, Michael Tomaszowski, **Heiko Fittler**, Andreas Christmann, Olga Avrutina, Harald Kolmar, Combinatorial optimization of cystine-knot peptides towards high-affinity inhibitors of human matriptase-1, PLoS One **2013**, 8, e76956

Vanessa Siegmund, Stefan Schmelz, Stephan Dickgiesser, Jan Beck, Aileen Ebenig, David Fiebig, **Heiko Fittler**, Holm Frauendorf, Birgit Piater, Ulrich Betz, Olga Avrutina, Andrea Scrima, Hans-Lothar Fuchsbauer, Harald Kolmar, Locked by design: A conformationally constrained transglutaminase tag enables efficient site-specific conjugation. *Angewandte Chemie* **2015**, DOI: 10.1002/anie.201504851.

2.1. Between two worlds: a comparative study on in vitro and in silico inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH

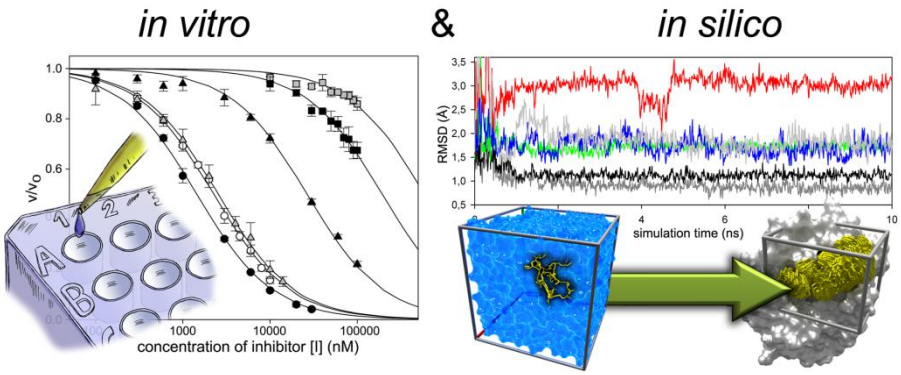
Authors:

Olga Avrutina, Heiko Fittler, Bernhard Glotzbach, Harald Kolmar, Martin Empting

Bibliographic Data:

Organic & Biomolecular Chemistry,
Volume 10, Issue 38, Page 7753-7762, 12 September 2012.
DOI: 10.1039/C3OB27469A
First published online: 6 August 2012.

Graphical Abstract:



Contributions by H. Fittler

- Synthesis of two compounds
- Performed all matriptase-inhibition assays for all compounds

Organic & Biomolecular Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 7753

www.rsc.org/obc

PAPER

Between two worlds: a comparative study on *in vitro* and *in silico* inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH†

Olga Avrutina, Heiko Fittler, Bernhard Glotzbach, Harald Kolmar and Martin Empting*

Received 18th June 2012, Accepted 3rd August 2012
DOI: 10.1039/c2ob26162f

A comparative study on *in vitro* and *in silico* inhibition of trypsin and matriptase by derivatives of the sunflower trypsin inhibitor-1 at near physiological pH is reported. Besides wild-type bicyclic SFTI-1, monocyclic variants possessing native cystine as well as redox-stable triazolyl side-chain macrocyclization motifs were studied for the first time in matriptase inhibition assays. Interestingly, monocyclic SFTI-1[*1,14*] demonstrated higher potency against this pharmacologically relevant protease compared to its bicyclic counterpart. Structural analysis of binding/inhibition of investigated SFTI-1 derivatives was performed using a combination of molecular dynamics simulations and docking experiments. *In silico* data were in good accordance with *in vitro* results, indicating the importance of the terminal inhibitor regions for the affinity towards matriptase. Presented work gives new perspectives for the optimization of the SFTI-1 framework towards *in vivo* applications.

Introduction

Very often interesting scientific findings arise from unexpected or even counterintuitive results. Recently, a comparison between the crystal structures of the sunflower trypsin inhibitor-1 (SFTI-1), a bicyclic tetradecapeptide, in complex with two closely related serine proteases revealed a surprising biochemical peculiarity.[1,2] By analyzing the binding geometry of SFTI-1 matriptase and trypsin aggregates, Yuan *et al.* confirmed a high similarity between both structures.[1,2] Though the homologous interaction profile of these serine proteases with SFTI-1 was expected, it appeared rather astonishing that this smallest naturally occurring Bowman–Birk Inhibitor (BBI)[3,4] favors trypsin over matriptase in retrospect.[1–3,5] Under the basic conditions (pH 8.2–8.5) commonly applied to *in vitro* inhibition assays for these enzymes, the electrostatic surface potential around the active site of matriptase shows, compared to the environment around the catalytic center of trypsin, a significantly higher negative polarization (Fig. 1). This provides a better overall charge complementarity with SFTI-1. It has been proposed that this anticipated enthalpic preference is overcompensated by an entropic penalty predominantly originating from the secondary binding loop of SFTI-1.[1] This might explain the surprising difference between its inhibitory activity against trypsin (K_i 0.1 nM) and matriptase (K_i 1–100 nM).[1–3,5]

However, selective binding and/or inhibition of matriptase under physiological conditions is of great pharmacological interest as this type II transmembrane serine protease (TTSP) is involved in a variety of pathological processes like progression and metastasis of epithelial tumors, as well as osteoarthritis and atherosclerosis.[6–10] A number of potent inhibitors of this serine protease based either on small organic molecules or large antibody fragments and exhibiting single digit nanomolar or even picomolar inhibition constants, respectively, have already been reported.[11,12] Nevertheless, attempts to tune the peptidic SFTI-1 framework towards matriptase resulted in inhibitors with enhanced selectivity but dramatically decreased potency.[5] Besides monocyclic variants lacking the disulfide connectivity and possessing substantially worsened kinetic properties, matriptase inhibition by side-chain cyclized SFTI-1 derivatives having free N- and C-termini has not been reported to date.

Recently, we developed potent inhibitors of trypsin based on monocyclic SFTI-1 derivatives containing redox-stable triazole bridges of different length and shape.[13] The impact of the macrocyclization motif in their bioactivity was demonstrated as well.[13] Herein, we report an unexpected change of the relative potency within the set of matriptase inhibitors **1–6** (Fig. 2) and evaluate

Clemens-Schöpf Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Petersenstr. 22, 64287 Darmstadt, Germany. E-mail: Empting@Biochemie-TUD.de

†Electronic supplementary information (ESI) available: Detailed synthetic procedures of compounds **1** and **3**, experimental details of enzyme assays, Fig. S1–S9 (SDS-Page analysis of inclusion body production, refolding/autoactivation, and purification of recombinant matriptase, as well as RP-HPLC traces, ESI-MS spectra, and dose–response curves against trypsin for previously uncharacterized inhibitors **1** and **3**), and Tables S1–S7 (determination of K_i values using alternative equations and the full set of used force field parameters). See DOI: 10.1039/c2ob26162f

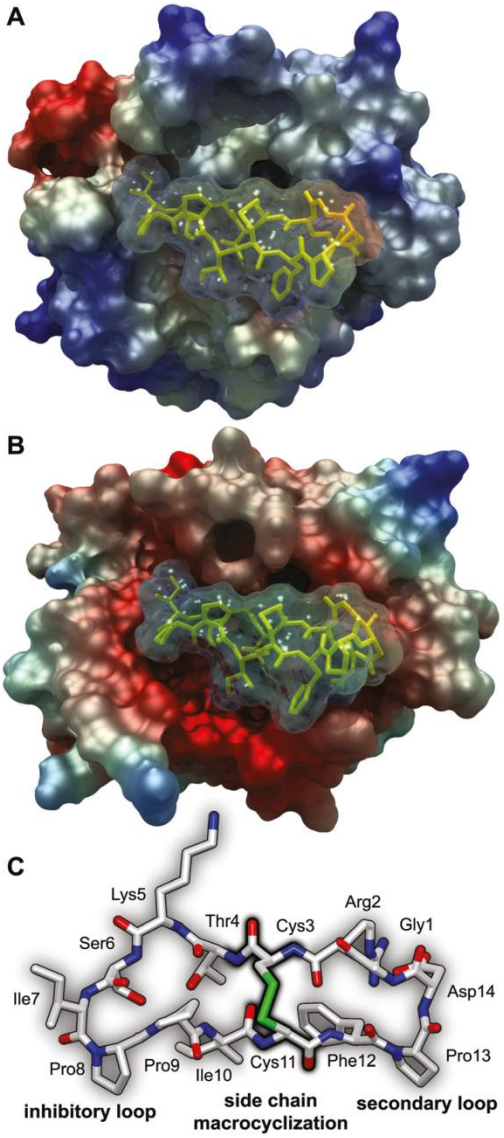


Fig. 1 Molecular surfaces of bicyclic SFTI-1 (transparent surface and yellow sticks) in complex with (A) trypsin (PDB ID: 1SFI) and (B) matriptase (PDB ID: 3P8F). Color gradients (red to blue) indicate electrostatic surface potentials (negative to positive) at pH 8.5 calculated *via* the Particle Mesh Ewald method with a maximum value of 500 kJ mol^{-1}.[14] (C) Stick representation of SFTI-1 structure. Blue: nitrogen, red: oxygen, green: sulfur, white: carbon, hydrogen omitted for clarity.

this finding on a structural basis using a two-step *in silico* method.

Recently, Swedberg and coworkers have used molecular dynamics (MD) simulations to predict beneficial amino acid exchanges for SFTI-1-derived inhibitors of kallikrein-related peptidase 4 (KLK4).[15] Our computational approach combines an MD calculation with subsequent semi-flexible and rigid docking experiments using a customized AMBER-derived force field (YASARA2) with embedded parameters for the calculation of non-natural triazolyl moieties within peptides.[16–18] This procedure enabled us to investigate the structural characteristics of

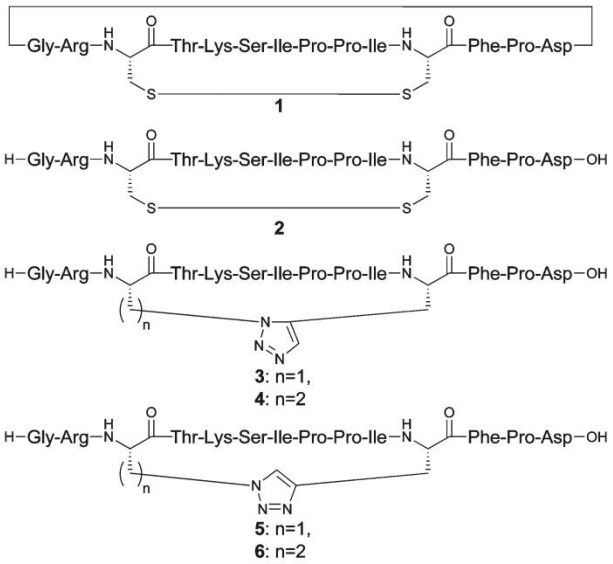


Fig. 2 Structures of investigated peptides and peptidomimetics. Bicyclic SFTI-1 (**1**), monocyclic SFTI-1[*1*,*14*] (**2**), monocyclic peptidomimetics containing either a 1,5-disubstituted 1,2,3-triazole (**3**, **4**) or a 1,4-disubstituted 1,2,3-triazole (**5**, **6**) as the macrocyclization motif.

each peptidic ligand in solution and upon binding to both trypsin and matriptase. *In silico* binding affinities were calculated and compared with the *in vitro* results. Thus, the role of the secondary loop as well as the impact of different macrocyclization motifs in matriptase inhibition of variants **1–6** were elucidated.

Results and discussion

In vitro inhibition of trypsin and matriptase

Trypsin is a prominent digestive protease commonly found in the small intestine of vertebrates[19] and operating most effectively under the mild basic conditions of its native environment (pH ~ 8).[20] Membrane-anchored matriptase, on the other hand, is usually expressed in epithelial cells and regulates cellular adhesion and growth by hydrolyzing proteins bound to cell surfaces or present in the extracellular matrix.[21] Thus, it functions in an environment controlled by homeostasis in healthy tissues (pH 7.4). Furthermore, tumors in general are known to generate hypoxic regions leading even to acidic microenvironments.[22–24] However, cancer-related matriptase possesses a pH optimum of about 9, therefore, the majority of reported inhibition assays have been carried out at a pH range between 8 and 9.[3,5,11,12,25,26]

We evaluated the potency of synthesized inhibitors through the determination of the fractional activity of the respective enzyme (v/v_0) in the presence of an inhibitor at various concentrations [I]. A number of mathematical methods are commonly used for the calculation of apparent inhibition constants K_i^{app} or IC_{50} values (concentration of the inhibitor to achieve a half-maximal degree of inhibition) for competitive inhibitors like SFTI-1.[27–32]

Therefore, we used the following equations for the non-linear fit of observed dose–response curves ([E] meaning active concentration of the enzyme).

$$\frac{v}{v_0} = \frac{1}{1 + \frac{[I]}{IC_{50}}} \quad (1)$$

$$\frac{v}{v_0} = \frac{([E] - [I] - K_i^{app}) + \sqrt{([I] + K_i^{app} - [E])^2 + 4K_i^{app}[E]}}{2[E]} \quad (2)$$

$$\frac{v}{v_0} = 1 - \frac{([E] - [I] + K_i^{app}) - \sqrt{([E] + [I] + K_i^{app})^2 - 4[E][I]}}{2[E]} \quad (3)$$

In the case of reversible competitive inhibition IC_{50} values calculated using eqn (1) can be converted into the substrate-independent K_i using eqn (4) ([S] is the concentration of the chromogenic substrate used within the enzyme assay and K_M is the corresponding Michaelis–Menten constant).[27]

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_M}} \quad (4)$$

In the case of tight-binding inhibition, eqn (5) has to be used instead of eqn (4).[28]

$$K_i = \frac{IC_{50} - \frac{E}{2}}{1 + \frac{[S]}{K_M}} \quad (5)$$

Eqn (2) and (3) are both adaptations of the Morrison equation for the determination of the substrate-dependent (apparent) inhibition constants for tight-binding inhibitors.[29–32] They can be transformed into each other *via* simple arithmetic conversions (see ESI†). Substrate-independent inhibition constant K_i can be calculated from K_i^{app} through eqn (6).[30]

$$K_i = \frac{K_1^{app}}{1 + \frac{[S]}{K_M}} \quad (6)$$

However, all used methods yielded essentially the same K_i values (see ESI†). For the sake of comparability with our previous studies, the results generated *via* eqn (3) and (6) are summarized in Table 1.

To investigate the effect of near physiological pH on matriptase inhibition by SFTI-1 (**1**) and SFTI-1[*1*,*14*] (**2**), we determined K_i values for both compounds at pH 7.6 and 8.5 (Fig. 3A, Table 1). A significant decrease in inhibitory activity for **1** and **2** at lower pH was observed. However, the detected inhibitory potency of bicyclic SFTI-1 under more basic conditions was in accordance with reported data (0.1 μM – Li *et al.*, 0.15 μM – this work).[5] Interestingly, the monocyclic variant SFTI-1[*1*,*14*] (**2**) possessing only a disulfide bond as the macrocyclization motif showed a better inhibition of matriptase compared to its backbone-cyclized counterpart **1**. To our knowledge, this unexpected finding has not been reported to date.

In the case of trypsin, the highly constrained bicyclic wild-type peptide (**1**) was the most potent inhibitor within the set of

Table 1 Determined K_i for compounds **1–6**

Entry	Trypsin/nM pH 7.6	Matriptase/μM pH 7.6	Matriptase/μM pH 8.5
1	0.07 ± 0.01	1.1 ± 0.14	0.15 ± 0.03
2	0.21 ± 0.03[a]	0.7 ± 0.09	0.1 ± 0.02
3	5.1 ± 1.0	1.24 ± 0.16	n.d.
4	0.34 ± 0.05[a]	12.9 ± 1.6	n.d.
5	273 ± 51[a]	285 ± 37	n.d.
6	106 ± 18[a]	94.1 ± 12	n.d.

All experiments were performed in triplicate. n.d. means not determined.[a] Values calculated from previously reported K_i^{app}.[13]

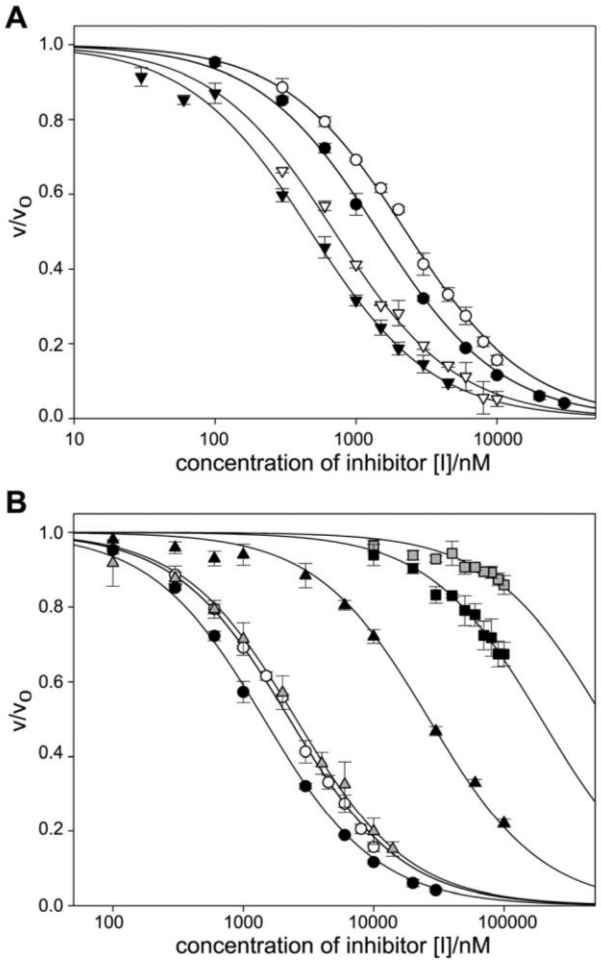


Fig. 3 Dose–response curves for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. (A) Comparison of bicyclic SFTI-1 (**1**) with monocyclic SFTI-1[*1*,*14*] (**2**) at pH 7.6 (white and black circles, respectively) and 8.5 (white and black triangles, respectively). (B) Comparison of disulfide-bridged inhibitors **1** (white circles) and **2** (black circles) with triazole-bridged peptidomimetics **3** (grey triangles), **4** (black triangles), **5** (grey squares), and **6** (black squares) at pH 7.6. Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

tested compounds **1–6**. Each modification of the SFTI-1 scaffold resulted in the loss of bioactivity against trypsin (Table 1). Nevertheless, the decrease in potency was quite marginal for open-chain variants **2** and **4**. The substitution pattern within the macrocyclization motif of peptidomimetics **3–6** had a dramatic influence on their bioactivity.[13] Inhibitors **3** and **4** containing a 1,5-disubstituted 1,2,3-triazole were significantly more potent compared to variants **5** and **6** possessing the 1,4-isomer.

However, novel compound **3** lacking one methylene unit within the side chain of the third residue (Fig. 2) was not as active as variant **4**.

To study the effect of the described SFTI-1 framework modifications on matriptase inhibition, we determined K_i values for all compounds at pH 7.6 (Fig. 3B, Table 1). Similarly to trypsin inhibition, peptidomimetics **3** and **4** showed better performance than **5** and **6** in matriptase assays as well. Surprisingly, comparison of variants **3** and **4** showed that **3** was more active against matriptase, while **4** was the better trypsin inhibitor.

The collected *in vitro* data shown in Table 1 emphasize that the native SFTI-1 scaffold has a general preference for trypsin over matriptase, especially at near physiological pH. The discrimination between trypsin and matriptase upon changing the target enzyme must have intrinsic structural reasons. Moreover, the unexpected inversion of relative affinities (**1** *versus* **2**, and **3** *versus* **4**) has to be explained.

MD simulation of SFTI-1 (1), SFTI-1[*1,14*] (2) and peptidomimetic compounds 3–6

To elucidate in atomic detail the impact of every single modification, we simulated compounds **1–6** as shown in Fig. 4 (see also the Experimental section).

A significant amount of structural data on SFTI-1 in solution as well as in complex with relevant proteases is available in the protein data base.[1,2,33] Thus, for the simulation of monocyclic variants **3–6** initial coordinates were derived from the reported ones (1JBN). However, preliminary studies revealed that the extensive hydrogen bond network of the SFTI-1 framework is able to compensate the steric strain on the peptide backbone generated by rigid, planar triazole bridges. This resulted in unnaturally bent heterocyclic structures within compounds **5** and **6** (ESI†). Thus, we decreased the scaling factor for non-bonded interactions and applied increased temperature during the initial phase of simulation (Fig. 4). This procedure allowed respective structures to leave the local minimum of the template coordinates. After incremental rescaling of parameters within 1 ns and an additional relaxation period of 8 ns no significant fluctuations in backbone root mean square deviations (RMSD) were observed (Fig. 5A). Hence, stable trajectories were analyzed in detail revealing fundamental structural differences between compounds **1–6** (Fig. 5B, C, and Table 2). The most dramatic effect on the peptide structure is caused by the 1,4-disubstituted 1,2,3-triazole bridge within variant **5**. Fig. 5B clearly illustrates a significant distortion of the inhibitory loop compared to crystal structures of wild-type SFTI-1 in complex with trypsin and matriptase. A similar, although not so dramatic, effect was observed for compound **6** which contained the 1,4-substitution pattern as well.

The applied simulation procedure resulted in significantly larger distances between the α-carbons of residues 3 and 11 for both compounds (**5** and **6**) compared to our previous study where only an energy minimization step was used.[13] Thus, the two opposing β-strands within the respective peptide frameworks are more remote from each other than within the structures of inhibitors **1–4**. Consequently, the possibility to form intramolecular hydrogen bonds is reduced for peptides **5** and **6** (Table 2). Since the hydrogen bond network is a structural prerequisite for the stability of the BBI loop conformation and, thus, biological activity,[15] these findings could explain the observed loss of potency for SFTI-1 variants possessing the 1,4-substitution pattern within their macrocyclization motif.

Bicyclic wild type **1** and monocyclic compound **2**, as well as 1,5-disubstituted 1,2,3-triazole containing peptidomimetics **3** and **4**, did not show a significant disorder of the inhibitor loop (Fig. 5B). Nevertheless, considerable structural changes within the N- and C-terminal regions of compounds **2–6** compared to the secondary loop of SFTI-1 occurred (Fig. 5C). This result, however, was highly anticipated as the conformational constraints caused by the additional backbone macrocyclization of peptide **1** are missing within the open-chain variants. The structural deviation of residues 1, 2 and 12–14 within highly potent trypsin inhibitor **4** was the most prominent one, while the values for compounds **3**, **5**, and **6** were in the range of monocyclic SFTI-1[1,14] (Fig. 5C). Considering the decreased relative activity of **4** against matriptase, our structural data support the notion that the N- and C-terminal regions are responsible for enzyme specificity rather than for the general capability to inhibit trypsin-like proteases. Thus, the observed structural changes and increased flexibility of these backbone segments in open-chain variants most likely are the reason for the improved activity of monocyclic compound **2** compared to bicyclic **1** against matriptase.

Docking of SFTI-1 (1), SFTI-1[*1,14*] (2) and peptidomimetic compounds 3–6 to trypsin and matriptase

To investigate the impact of described framework modifications in the binding geometries of inhibitor–protease complexes, docking experiments were performed (see the Experimental section, Fig. 4). The collected data are summarized in Table 3.

As all of the investigated compounds **1–6** inhibit trypsin and matriptase in a concentration-dependent manner (Fig. 3 and Table 1), we rationalized that interaction of open-chain variant **2** and peptidomimetics **3–6** with these proteases must follow a similar binding geometry to wild-type SFTI-1. This postulation is supported by recently solved inhibitor–trypsin complex structures of monocyclic SFTI-based peptidomimetics.[34] Thus, in order to exhibit observed biological activities, each compound must penetrate the S1 site of a respective protease with the side chain of Lys5. Consequently, we screened the set of generated inhibitor–protease complexes for structures that fulfilled this requirement. The ratio of counted aggregates matching the expected binding geometry to the number of all calculated structures (100 each) is given as P_{hit} (Table 3). Interestingly, in our docking experiments this value was significantly higher for trypsin than for matriptase. Furthermore, all potent inhibitors

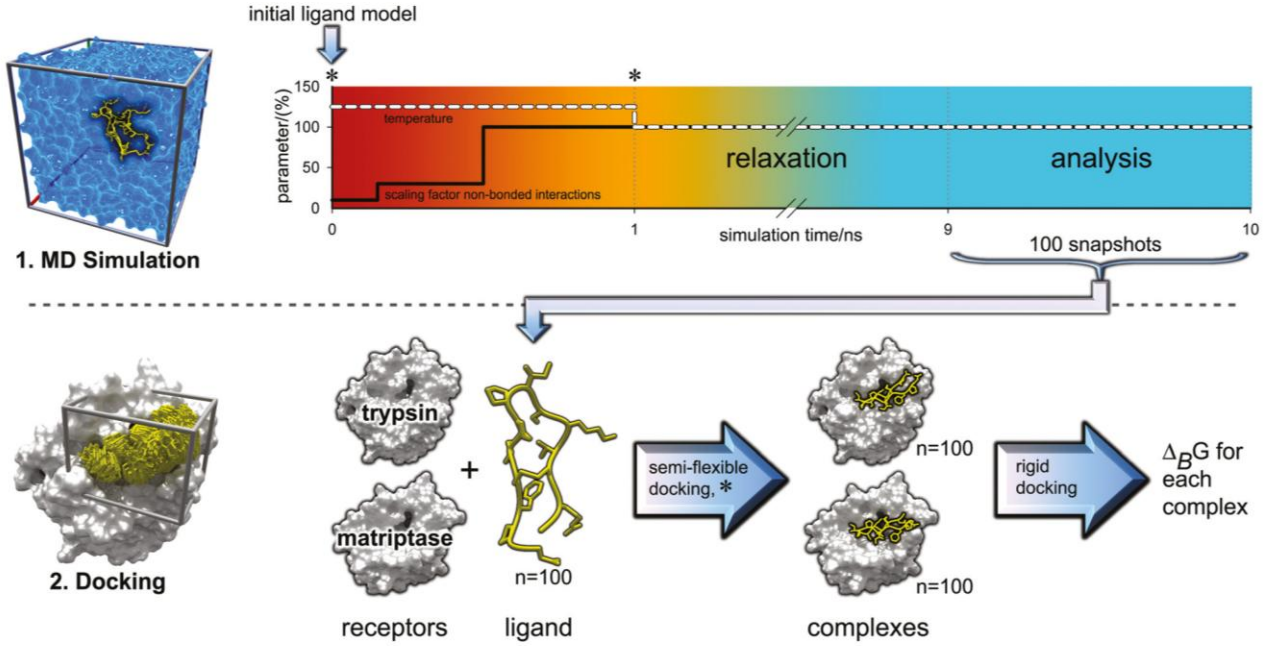


Fig. 4 Design of the two-step *in silico* experiment. Illustrative images (left) show a simulation cell containing water and a peptidic ligand (top left), as well as a number of ligand poses generated in a local docking experiment restricted by a simulation cell around the active site of the receptor (bottom left). The MD simulation procedure (top) is depicted as time-lapsed absolute temperature (dashed, white) and scaling factor curves for non-bonded interactions (black, solid). Both parameters are given in percent to 298 K and 1, respectively. The color gradient (red to blue) qualitatively illustrates the resulting stress applied to the simulated structure; red means harsh and blue means normalized conditions. The 8 ns period for relaxation is shortened. Snapshots between nanosecond nine and ten are analyzed using a docking procedure (bottom). Free energies of binding $\Delta_B G$ are calculated by sequential semi-flexible and rigid docking steps. Receptor molecules trypsin and matriptase are shown as white surfaces. The ligand is depicted as yellow tubes and sticks. The star symbol indicates an energy minimization step.

1–**4** consistently showed a higher probability P_{hit} to interact with both proteases in a mode similar to the reported complexes of SFTI-1, compared to weak binders **5** and **6**.

We also investigated the capability of compounds **1**–**6** to form intermolecular hydrogen bonds with each of the target proteases.

Thus, the number of corresponding interactions was counted for each complex with the ligand docked following the described binding mode. As expected, peptidomimetic **5** showed the lowest number of hydrogen bonds within the set of examined structures. However, the amount of measured intermolecular proton donor–acceptor interactions of compound **6** with trypsin and matriptase was in the same range as for inhibitors **1**–**4**.

The analysis of the structural and thermodynamic properties of the generated docking results was of particular importance. The applied procedure enabled us to determine the free energy of binding $\Delta_B G$ for each ligand–receptor complex (Table 3). Calculated structures of compounds **1**–**6** bound to matriptase and possessing the highest absolute value of $\Delta_B G$ are shown in Fig. 6. These models exhibited the most favorable interaction profiles observed, therefore were the best predicted structures for the actual binding geometry within our *in silico* experiment. Complexes **1** and **2** were found to interact with matriptase in a very similar manner compared to the reported crystal structure (Fig. 6A, B). Due to the highly distorted BBI loop of peptidomimetic **5**, this variant showed the most deviating docking geometry (Fig. 6E). The N- and C-terminal regions were drastically displaced, considerably reducing the possibility for attractive interactions. Indeed, only a moderate increase of the backbone RMSD values was observed for compounds **3**, **4**, and **6**. Similar data were extracted from docking experiments with trypsin (Fig. S9, ESI†). However, a more prominent deviation between predicted binding geometries and the crystal structure was measured for the residues corresponding to the secondary loop of SFTI-1. The stability of the investigated aggregates was evaluated by averaging calculated $\Delta_B G$ for every structure with expected binding geometry resulting in $\Delta_B G^{dock}$ (Table 3).

In general, the calculated mean values indicated a significantly more favorable interaction between matriptase and inhibitors **1**–**6** compared to respective complexes with trypsin. This finding corroborated the expected influence of charge complementarity between the SFTI-1 scaffold and the surface around the active site of matriptase on overall complex stability.

The most prominent deviations upon docking to trypsin are observed in the N- and C-terminal regions of compounds **1**–**6**. This indicates less conformational constraints compared to the respective matriptase complexes. Thus, an increased flexibility of bound SFTI ligands is provided, while the amount of attractive electrostatic interactions is reduced. A tighter, but also more constrained binding of SFTI-1 to matriptase had already been postulated by Yuan *et al.* and was corroborated by our experiments.[1]

Surprisingly, the ability to bind similarly to the reported crystal structures did not necessarily lead to high values of mean $\Delta_B G^{dock}$. Thus, for variant **6** the best docking complexes were characterized by rather low backbone RMSD values (Fig. 6F and

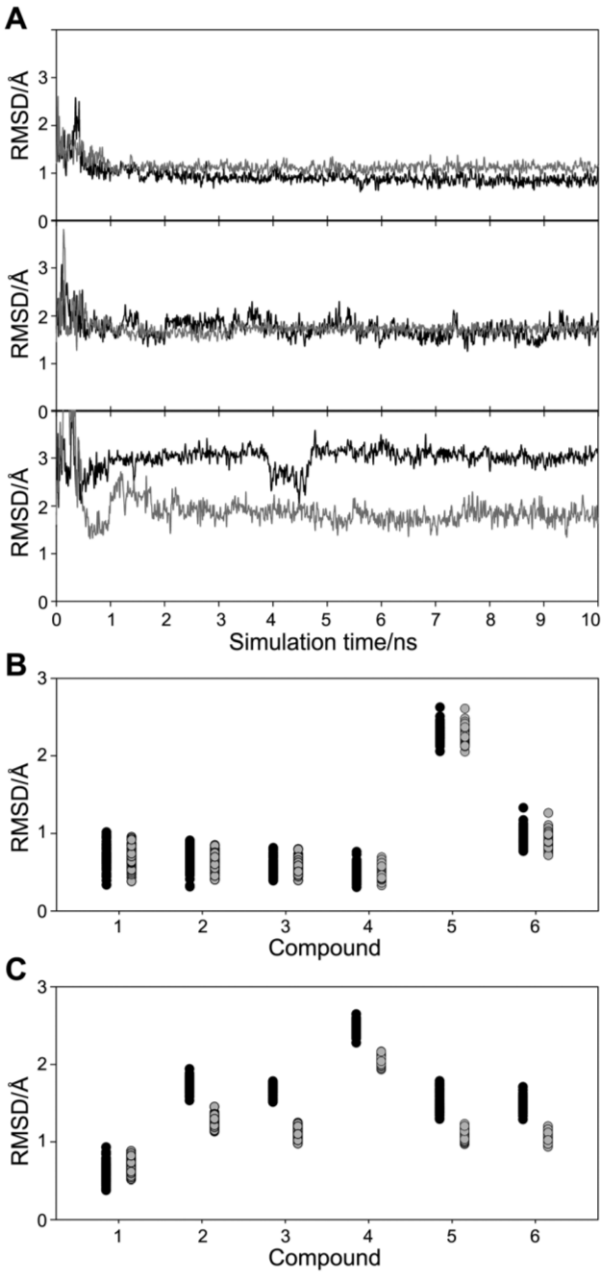


Fig. 5 (A) Backbone RMSD values of **1** (top, black), **2** (top, grey), **3** (middle, black), **4** (middle, grey), **5** (bottom, black), and **6** (bottom, grey) compared to SFTI-1 solution structure 1JBL plotted against the simulation time. Dot plots of backbone RMSD values of the inhibitory loop (B) and the secondary loop/N- and C-terminal regions (C) of compounds **1–6** compared to SFTI-1 in complexes with trypsin (black) and matriptase (grey) showing 100 values determined every 10 ps during the last nanosecond of simulation.

S9F). Nevertheless, more distorted structures with significantly lower affinities were also observed. This heterogeneity of interaction profiles was reflected by the highest errors (standard deviations) of corresponding $\Delta_B G^{dock}$ within the set of investigated compounds (Table 3). This probably originates from the described moderate distortion of the inhibitory loop during simulation (Fig. 5B). Nevertheless, the applied two-step docking procedure allowed a number of the *in silico* samples of peptidomimetic **6** to reconfigure the backbone conformation to the original BBI geometry upon interaction with the target enzymes.

Table 2 Data collected from the simulation experiments for compounds **1–6**

Entry	Backbone RMSD[a,b]/Å	Distance Cα3–Cα11[b]/Å	Intramol. H-bonds[b,c]
1	0.87 ± 0.09	3.93 ± 0.08	6.08 ± 1.08
2	1.13 ± 0.07	3.95 ± 0.07	6.12 ± 1.14
3	1.72 ± 0.15	3.90 ± 0.06	5.85 ± 1.15
4	1.74 ± 0.06	4.29 ± 0.10	5.67 ± 0.96
5	3.02 ± 0.10	5.66 ± 0.12	2.63 ± 0.91
6	1.79 ± 0.15	5.16 ± 0.14	4.36 ± 0.82

[a] Compared to solution structure 1JBL. [b] Values are arithmetical means of 100 structures taken from the last nanosecond of simulation. Errors are given as standard deviation. [c] Measured using a threshold of 6.25 kJ mol^{-1}.

In our docking experiments, compound **4** showed average values of $\Delta_B G^{dock}$ very similar to **6** (Table 3). In this case, the pronounced deviation within the N- and C-terminal regions can be considered as the reason for the reduced complex stability (Fig. 5C). Thus, within the predicted inhibitor **4**–matriptase complex the aromatic side chain of Phe12 is positioned considerably different to inhibitors **1–3** (Fig. 6D). In the corresponding crystal structure this residue is highly constrained, being flanked by two benzyl groups of Phe97 and Phe99 of the protease molecule.[1] Thus, affinity of peptidomimetic **4** towards matriptase might be reduced due to a loss of favorable hydrophobic interactions within this region. Upon docking to trypsin, the C-terminal residues were even more displaced (Fig. S9D†). Nevertheless, the observed loss of free binding energy was still only of moderate magnitude for compound **4**. Interestingly, this inhibitor demonstrated the highest probability to interact with both proteases in the BBI geometry within the set of peptidomimetics **3–6** (Table 3).

Compound **3** showed significantly lower P_{hit} values compared to **4**. However, if bound in the BBI modality, **3** formed far more stable complexes. The determined absolute values of $\Delta_B G^{dock}$ were even higher than those of monocyclic **2**. As, compared to **4**, peptidomimetic **3** lacks one methylene unit within the macrocyclization motif (Fig. 2), this structure possesses less degrees of conformational freedom. As a result, the proximity of α-carbons between residues 3 and 11 in compound **3** matches the corresponding distance within the wild-type peptides **1** and **2** almost perfectly (Table 2). Nevertheless, the discussed structural constraints also caused the loss of backbone flexibility. Thus, compound **3** is held within a defined range of conformations displaying high stabilities of complexes with both proteases, once the inhibitor is positioned accordingly. However, the applied docking algorithm did not always succeed to fit compound **3** into the BBI geometry. This indicates that the latitude between perfect fit and steric clashes for SFTI-1 variants in complex with matriptase is rather limited.

A similar effect was observed for wild type **1**. This very constrained bicyclic inhibitor consistently showed the highest

Table 3 Data collected from the docking experiments for compounds **1–6** against trypsin and matriptase and corresponding *in vitro* free energies of binding $\Delta_B G^{exp}$

	Trypsin					Matriptase				
Entry	P_{hit}/ (%)	Intermol. H-bonds[a,b]	$\Delta_B G^{dock}$ [a]/ (kJ mol^{-1})	$\Delta_B G^{calc}$ [c]/ (kJ mol^{-1})	$\Delta_B G^{exp}$ [d]/ (kJ mol^{-1})	P_{hit}/ (%)	Intermol. H-bonds[a]	$\Delta_B G^{dock}$ [a]/ (kJ mol^{-1})	$\Delta_B G^{calc}$ [c]/ (kJ mol^{-1})	$\Delta_B G^{exp}$ [d]/ (kJ mol^{-1})
1	96	10.3 ± 1.9	61.3 ± 6.3	58.9 ± 6.1	58.0 ± 0.4	51	9.2 ± 1.8	71.0 ± 8.1	36.2 ± 4.1	34.0 ± 0.3
2	99	11.3 ± 1.5	51.6 ± 4.9	51.1 ± 4.8	55.2 ± 0.4	67	10.1 ± 2.1	63.2 ± 4.9	42.3 ± 5.8	35.1 ± 0.3
3	87	10.9 ± 1.4	54.8 ± 6.9	47.7 ± 6.0	47.3 ± 0.5	51	11.0 ± 1.6	69.0 ± 6.9	35.2 ± 4.4	33.7 ± 0.3
4	100	10.0 ± 1.7	48.6 ± 5.4	48.6 ± 5.4	54.0 ± 0.3	61	9.1 ± 1.9	56.4 ± 5.4	34.4 ± 7.1	27.9 ± 0.3
5	69	8.9 ± 1.4	33.6 ± 3.9	23.2 ± 2.7	37.5 ± 0.5	47	8.4 ± 1.5	42.6 ± 3.9	20.0 ± 3.0	20.2 ± 0.3
6	81	9.8 ± 1.6	43.8 ± 9.8	35.4 ± 8.0	39.8 ± 0.4	43	9.3 ± 2.2	56.1 ± 9.8	24.1 ± 4.8	23.0 ± 0.3

[a] Values are arithmetical means of all inhibitor–protease complexes following BBI modality and errors are standard deviations. [b] Measured using a threshold of 6.25 kJ mol^{-1}. [c] Calculated using $\Delta_B G^{calc} = P_{hit} \times \Delta_B G^{dock}$. [d] Calculated using eqn (7). Errors of $\Delta_B G^{exp}$ were calculated by propagation of errors (ESI†).

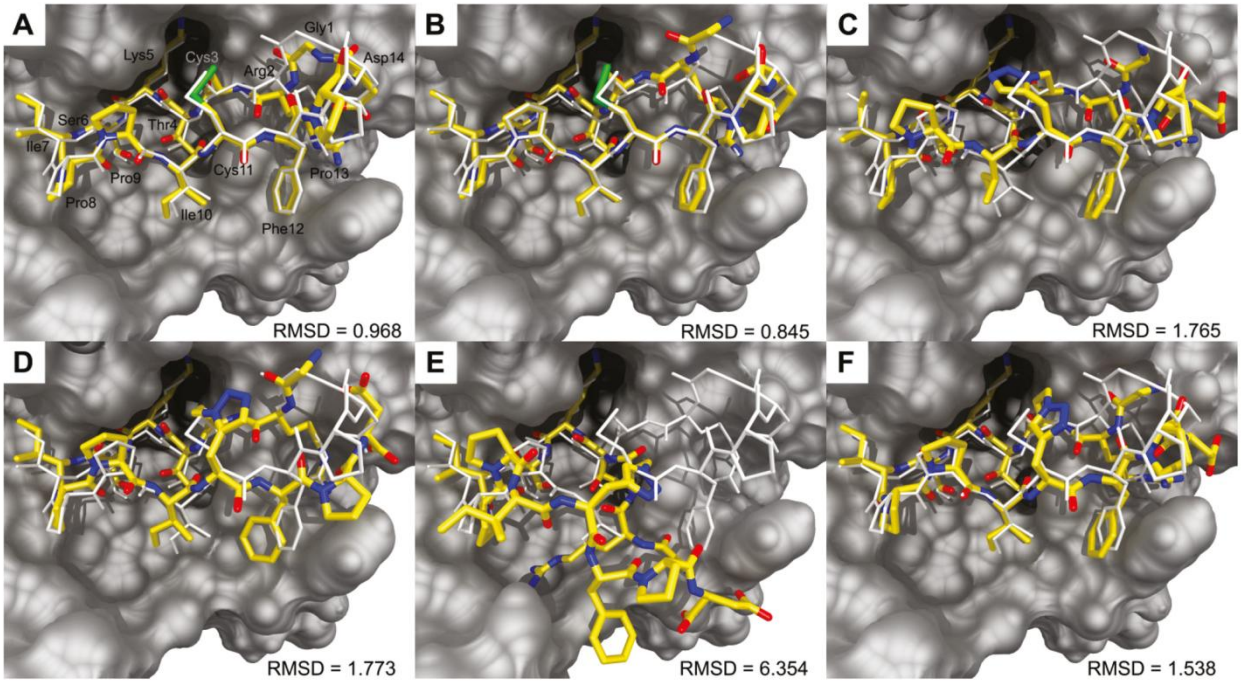


Fig. 6 Predicted structures of compounds **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), and **6** (F) in complex with matriptase (grey surface) as an overlay with reported crystal structure 3P8F (white sticks).[1] Blue: nitrogen, green: sulfur, red: oxygen, yellow: carbon, hydrogen is omitted for clarity. Measured RMSD values for inhibitor backbones compared to 3P8F are given in Å.

values of $\Delta_B G^{dock}$ upon binding to both enzymes. Nevertheless, the probability of **1** to actually interact under the mode revealed by crystal structures was significantly lower than for compounds **2** and **4**. Thus, flexibility within the region of the secondary loop might be crucial for the inhibitor to be effectively transferred from solution (simulation) into the active site of the proteases (docking). This effect is far more pronounced in the case of matriptase, probably due to the mentioned reduced space around residue 12 of inhibitors **1–6**.

Monocyclic inhibitor **2** seems to provide a good compromise between flexibility and proper backbone arrangement. Hence, it forms very stable complexes with both proteases at a high probability.

Comparison of *in vitro* and *in silico* data

To take into account the reduced probabilities for SFTI-1 derivatives to bind to matriptase according to elucidated crystal structures, we directly weighted all $\Delta_B G^{dock}$ by corresponding P_{hit} to calculate *in silico* binding affinities $\Delta_B G^{calc}$ (Table 3). This procedure allowed evaluating relative potencies of the investigated compounds towards examined proteases. Additionally, experimental inhibition constants K_i were used to calculate free energies of binding $\Delta_B G^{exp}$ using eqn (7) (R meaning the universal gas constant and T the absolute temperature).

$$\Delta_B G^{exp} = -RT \ln K_i \quad (7)$$

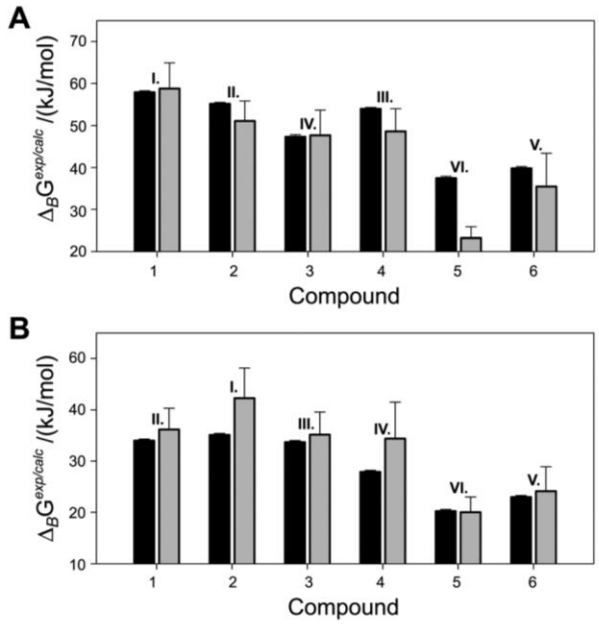


Fig. 7 Bar chart of *in vitro* $\Delta_B G^{exp}$ (black) and *in silico* $\Delta_B G^{calc}$ (grey) for complexes of compounds **1–6** with trypsin (A) and matriptase (B). Rankings of binding affinities for both *in vitro* and *in silico* experiments are indicated by attic numerals (best to worst: I.–VI.).

Thus, the collected *in silico* data were compared with corresponding *in vitro* results (Fig. 7).

The applied simulation/docking procedure enabled to fully reproduce the outcome of the enzyme assays qualitatively. Thus, ranking investigated inhibitors **1–6** according to their affinity ($\Delta_B G^{calc}$ or $\Delta_B G^{exp}$) resulted in the same relative order for both proteases (Fig. 7). Generally, absolute values of $\Delta_B G^{calc}$ for the inhibition of matriptase were higher compared to those determined by *in vitro* experiments. Since calculated binding affinities can only be regarded as rough estimates, both $\Delta_B G^{calc}$ and $\Delta_B G^{exp}$ are in astonishingly good agreement.

A similar result was achieved for trypsin inhibition. However, the most significant difference between *in vitro* and *in silico* data was observed for the weakest inhibitor **5** (Fig. 7A). Thus, it is probable that the applied docking procedure overestimates the influence of the massive distortion of the BBI loop in solution on trypsin affinity. The dramatic displacement of the peptide backbone may actually be attenuated through dynamic interactions of **5** with the surface of the protease. Nevertheless, the possibility for minor backbone readjustments was provided by energy minimization between the semi-flexible and the rigid docking steps. An additional time-intensive MD simulation procedure might be required to cover the dynamic nature of ligand–receptor interactions as well as the influence of solvent molecules on the investigated complex structures. Furthermore, effects of transient or permanent breakage and formation of covalent bonds have to be ignored within the classical force field approach. Such reactions might have a significant impact in the investigated system, as dynamic hydrolysis/recyclization equilibrium has been reported for trypsin inhibition by SFTI-1.[35]

Conclusion

Our *in vitro* data show that the *in vivo* application of the wild-type SFTI-1 as a therapeutic agent targeting matriptase is limited due to its drastically reduced potency at physiological pH. Nevertheless, as reported in previous works, the SFTI-1 framework provides an appropriate starting point for lead compound optimization towards inhibition of pharmacologically relevant proteases.[15,36–38] With the focus on matriptase as the target enzyme, we suggest the open-chain variant SFTI-1[*1,14*] as a promising scaffold rather than its bicyclic ancestor. Based on these findings, we developed SFTI-1[*1,14*] derivatives with significantly improved affinity and selectivity towards the addressed TTSP; these results will be published elsewhere.

Additionally, triazole-bridged matriptase inhibitor **3** provides the best kinetic properties within the set of peptidomimetics **3–6**. Its highly constrained side-chain macrocyclic motif might also be a valuable asset when redox stability is of particular importance within the pharmaceutical context.

Our *in silico* experiments provided a plausible explanation for the potencies observed *in vitro*. Despite the intrinsic limitations and simplifications of the force field approach, computed binding affinities $\Delta_B G^{calc}$ represented satisfactory qualitative estimates of values determined by enzyme assays. Structural analysis of the simulation and docking experiments indicated the importance of the terminal regions of SFTI-1 derivatives upon matriptase binding/inhibition. Higher conformational constraints lead to better interaction profiles while lowering the overall probability to bind matriptase following the expected modality. Nevertheless, changes within the structure of the BBI macrocycle affected inhibitory potency in the most severe manner. However, further refinement of force field parameters for triazolyl moieties within peptidic molecules using additional X-ray and/or NMR experiments might be necessary for more precise calculations. Additionally, to confirm the predicted entropic effects on inhibitory potency, isothermal calorimetry could be conducted.[39]

Experimental

Chemicals

All Fmoc-protected amino acids, resins for solid phase peptide synthesis (SPPS), solvents, and other chemicals were purchased from Agilent (Varian), Bachem, Fischer Chemicals, Iris Biotech, Novabiochem, Roth and Sigma-Aldrich.

Peptide synthesis

Linear peptide precursors were assembled by automated microwave-assisted SPPS using the Fmoc strategy.[40,41]

Bicyclic SFTI-1 (**1**) was synthesized *via* solution-phase backbone macrocyclization of linear precursor H-Arg(Pbf)-Cys(Trt)-Thr(*t*Bu)-Lys(Boc)-Ser(*t*Bu)-Ile-Pro-Pro-Ile-Cys(Trt)-Phe-Pro-Asp-(*t*Bu)-Gly-OH using C-terminal PyBOP/HOBt activation followed by acidolytic side-chain deprotection, oxidative disulfide bond formation and chromatographic isolation.

Synthesis and characterization of monocyclic SFTI-1[*1,14*] (**2**) and triazole-containing peptidomimetics **4–6** have been

previously reported.[13] New peptidomimetic compound **3** possessing a 1,5-disubstituted 1,2,3-triazolyl side chain macrocyclization motif was synthesized using a similar approach.[13] First, the peptide resin containing unnatural amino acids azidoalanine (Aza) and propargylglycine (Pra) (Fmoc-Aza-Thr(*t*Bu)-Lys (Boc)-Ser(*t*Bu)-Ile-Pro-Pro-Ile-Pra-Phe-Pro-Asp(*t*Bu)-resin) was synthesized. Then, the triazole bridge was installed *via* on-support ruthenium(II)-catalyzed azide–alkyne cycloaddition (RuAAC) followed by coupling of N-terminal residues Gly1 and Arg2. Acidolytic cleavage and chromatographic isolation followed.

For experimental details and analytical data on bicyclic SFTI-1 (**1**) and new triazole-bridged compound **3** refer to ESI.†

Recombinant production of matriptase

BL21-CodonPlus (DE3)-RP competent *E. coli* cells (Stratagene) were transformed with expression vector pET42dest-His-hMatI (cd)596–855. This vector was constructed for overexpression of the zymogen sequence of the catalytic domain of matriptase as the following fusion: Met-Ser-Tyr-Tyr-His6-aa596–855. Inclusion body production, protein denaturation, purification and refolding/autoprotolytic activation were conducted according to reported procedures (see ESI, Fig. S1 and S2† for details).[11,25,42] Substrate-specific catalytic activity of the purified enzyme was verified through the hydrolysis of chromogenic substrate Boc-QAR-*p*NA and quantified *via* active-site titration using 4-methylumbelliferyl-*p*-guanidinobenzoate (MUGB) (see ESI, Fig. S3†).[43,44]

Inhibition assays

Apparent (substrate-dependent) inhibition constants K_i^{app} of SFTI-1-derivatives **1–6** against active-site-titrated bovine trypsin ([E] = 0.5 nM) and matriptase ([E] = 0.9 nM) were determined by recording kinetic curves of the enzymatic hydrolysis of chromogenic substrate Boc-QAR-*p*NA ([S] = 250 μM) at different concentrations of inhibitor [I] (pH 7.6 or pH 8.5) using a Tecan GENios microplate reader.[13] Calculated relative initial velocities of the residual proteolytic reaction (v/v_0) were plotted against [I]. Eqn (3) for tight-binding inhibitors was fitted by non-linear regression to the resulting dose–response curves using the Marquardt–Levenberg algorithm of Sigma Plot 11.[29]

Substrate-independent inhibition constants K_i were calculated from K_i^{app} and the Michaelis–Menten constants K_M for the corresponding enzyme–substrate complex using eqn (6).[25,34]

K_M for Boc-QAR-*p*NA in complex with matriptase at pH 7.6 and pH 8.5 were calculated by Lineweaver–Burk plots as 236.8 ± 56.1 μM and 66.6 ± 16.0 μM, respectively. K_M for trypsin as well as the method for the estimation of errors for each individual K_i have been recently reported.[34]

In the case of reversible competitive inhibition K_i equals the dissociation constant K_d of the inhibitor–enzyme complex.[45] Thus, the free energy of binding $\Delta_B G$ can be calculated from K_i using eqn (7).

Force field customization

All *in silico* experiments were performed with the YASARA structure package (YASARA Biosciences). For the precise handling of triazole-bridged peptides during *in silico* calculations, new topology entries and atom types for the involved carbon and nitrogen species have been implemented into the YASARA2 force field description file. Thus, amino acid residues with 1,4- and 1,5-disubstituted 1,2,3-triazoles in the side chain of compounds **3–6** were covered. Corresponding equilibrium bond lengths, bond angles and dihedrals have been deduced from reported crystal structures.[46,47] Values for force constants and scaling factors were adopted from similar aromatic and heteroaromatic structures (*e.g.* imidazole) within the original force field. The full set of parameters used to calculate disubstituted 1,2,3-triazoles is given in the ESI.†

Molecular dynamics simulation and docking

Solution NMR structures of SFTI-1 (PDB ID: 1JBL) and SFTI-1 [*1,14*] (PDB ID: 1JBN) were used as initial templates for compounds **1–6**.[33] Triazole bridges of peptidomimetics **3–6** were modelled into 1JBN to match the topology entry within the customized force field. Each compound was then subjected to a two-step simulation and docking procedure performed *via* a self-written command sequence (macro). The design of the *in silico* experiment is sketched in Fig. 4.

Following point charge assignment *via* semi-quantitative quantum mechanics,[48] an energy minimization was conducted with initial structures of compounds **1–6** in 0.9% (m/v) NaCl (aq.) at pH 7.6. To prevent structures from being trapped in the local minimum of the template coordinates, all compounds were then simulated at 373 K and the scaling factor for non-bonded interactions was reduced to 10%. This parameter was stepwise re-adjusted to 100% within 500 ps. After 1 ns of simulation, the temperature was set to 298 K and energy-minimized structures were simulated for another 9 ns. Trajectories (snapshots) were saved every 10 ps. Thus, 100 models for each individual compound **1–6** from the last nanosecond of the simulation were extracted and used for the docking experiment. Semi-flexible docking to the active sites of trypsin and matriptase with fixed ligand backbones and adjusted side-chain atoms of Lys5 was performed to identify the most probable binding geometry using the AutoDock 4 algorithm with 999 runs.[18] The highest ranked complex structure was subjected to an energy minimization with fixed receptor atoms followed by another energy minimization step with a fully unrestrained ligand–receptor complex. The resulting structures were then used for the rigid docking experiment (additional 999 runs) to compute virtual free energies of binding $\Delta_B G$ with AutoDock.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft through grant Ko 1390/9-1 and by BMBF. We thank Prof. Dr Kay Hamacher (AG Computational Biology & Simulation, Technische Universität Darmstadt) for discussion concerning *in silico* experiments and for reading the manuscript.

References

1 C. Yuan, L. Q. Chen, E. J. Meehan, N. Daly, D. J. Craik, M. D. Huang and J. C. Ngo, *BMC Struct. Biol.*, 2011, 11.
2 S. Luckett, R. S. Garcia, J. J. Barker, A. V. Konarev, P. R. Shewry, A. R. Clarke and R. L. Brady, *J. Mol. Biol.*, 1999, **290**, 525–533.
3 Y. Q. Long, S. L. Lee, C. Y. Lin, I. J. Enyedy, S. M. Wang, P. Li, R. B. Dickson and P. P. Roller, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 2515–2519.
4 M. L. J. Korsinczky, R. J. Clark and D. J. Craik, *Biochemistry*, 2005, **44**, 1145–1153.
5 P. Li, S. Jiang, S. L. Lee, C. Y. Lin, M. D. Johnson, R. B. Dickson, C. J. Michejda and P. P. Roller, *J. Med. Chem.*, 2007, **50**, 5976–5983.
6 Y. E. Shi, J. Torri, L. Yieh, A. Wellstein, M. E. Lippman and R. B. Dickson, *Cancer Res.*, 1993, **53**, 1409–1415.
7 K. Uhland, *Cell. Mol. Life Sci.*, 2006, **63**, 2968–2978.
8 K. List, *Future Oncol.*, 2009, **5**, 97–104.
9 J. M. Milner, A. Patel, R. K. Davidson, T. E. Swingler, A. Desilets, D. A. Young, E. B. Kelso, S. T. Donell, T. E. Cawston, I. M. Clark, W. R. Ferrell, R. Plevin, J. C. Lockhart, R. Leduc and A. D. Rowan, *Arthritis Rheum*, 2010, **62**, 1955–1966.
10 I. Seitz, S. Hess, H. Schulz, R. Eckl, G. Busch, H. P. Montens, R. Brandl, S. Seidl, A. Schomig and I. Ott, *Arterioscler., Thromb., Vasc. Biol.*, 2007, **27**, 769–775.
11 T. Steinmetzer, A. Schweinitz, A. Sturzebecher, D. Donnecke, K. Uhland, O. Schuster, P. Steinmetzer, F. Muller, R. Friedrich, M. E. Than, W. Bode and J. Sturzbecher, *J. Med. Chem.*, 2006, **49**, 4116–4126.
12 C. J. Farady, J. Sun, M. R. Darragh, S. M. Miller and C. S. Craik, *J. Mol. Biol.*, 2007, **369**, 1041–1051.
13 M. Empting, O. Avrutina, R. Meusinger, S. Fabritz, M. Reinwarth, M. Biesalski, S. Voigt, G. Buntkowsky and H. Kolmar, *Angew. Chem., Int. Ed.*, 2011, **50**, 5207–5211.
14 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
15 J. E. Swedberg, S. J. de Veer, K. C. Sit, C. F. Reboul, A. M. Buckle and J. M. Harris, *PLOS One*, 2011, 6.
16 E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker and K. Karplus, *Proteins: Struct., Funct., Bioinf.*, 2009, **77**, 114–122.
17 W. Brandt, T. Herberg and L. Wessjohann, *Biopolymers*, 2011, **96**, 651–668.
18 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639–1662.
19 W. R. Rypniewski, A. Perrakis, C. E. Vorgias and K. S. Wilson, *Protein Eng., Des. Sel.*, 1994, **7**, 57–64.
20 S. N. S. Murthy, J. Kostman and V. P. Dinoso, *Dig. Dis. Sci.*, 1980, **25**, 289–294.
21 T. H. Bugge, T. M. Antalis and Q. Y. Wu, *J. Biol. Chem.*, 2009, **284**, 23177–23181.
22 J. L. Wike-Hooley, J. Haveman and H. S. Reinhold, *Radiother. Oncol.*, 1984, **2**, 343–366.
23 I. F. Tannock and D. Rotin, *Cancer Res.*, 1989, **49**, 4373–4384.
24 P. Montcourrier, I. Silver, R. Farnoud, I. Bird and H. Rochefort, *Clin. Exp. Metastasis*, 1997, **15**, 382–392.
25 A. Desilets, J. M. Longpre, M. E. Beaulieu and R. Leduc, *FEBS Lett.*, 2006, **580**, 2227–2232.
26 F. Beliveau, A. Desilets and R. Leduc, *FEBS J.*, 2009, **276**, 2213–2226.
27 Y. Cheng and W. H. Prusoff, *Biochem. Pharmacol.*, 1973, **22**, 3099–3108.
28 R. A. Copeland, D. Lombardo, J. Giannaras and C. P. Decicco, *Bioorg. Med. Chem. Lett.*, 1995, **5**, 1947–1952.
29 J. F. Morrison, *Biochim. Biophys. Acta, Enzymol.*, 1969, **185**, 269–286.
30 J. W. Williams and J. F. Morrison, *Methods Enzymol.*, 1979, **63**, 437–467.
31 W. R. Greco and M. T. Hakala, *J. Biol. Chem.*, 1979, **254**, 12104–12109.
32 D. J. Murphy, *Anal. Biochem.*, 2004, **327**, 61–67.
33 M. L. J. Korsinczky, H. J. Schirra, K. J. Rosengren, J. West, B. A. Condie, L. Otvos, M. A. Anderson and D. J. Craik, *J. Mol. Biol.*, 2001, **311**, 579–591.
34 M. Tischler, D. Nasu, M. Empting, S. Schmelz, D. W. Heinz, P. Rottmann, H. Kolmar, G. Buntkowsky, D. Tietze and O. Avrutina, *Angew. Chem., Int. Ed.*, 2012, **51**, 3708–3712.
35 U. C. Marx, M. L. J. Korsinczky, H. J. Schirra, A. Jones, B. Condie, L. Otvos and D. J. Craik, *J. Biol. Chem.*, 2003, **278**, 21782–21789.
36 A. Lesner, A. Legowska, M. Wysocka and K. Rolka, *Curr. Pharm. Des.*, 2011, **17**, 4308–4317.
37 E. Zablotna, A. Jaskiewicz, A. Legowska, H. Miecznikowska, A. Lesner and K. Rolka, *J. Pept. Sci.*, 2007, **13**, 749–755.
38 J. D. McBride, H. N. M. Freeman and R. J. Leatherbarrow, *Eur. J. Biochem.*, 1999, **266**, 403–412.
39 S. Leavitt and E. Freire, *Curr. Opin. Struct. Biol.*, 2001, **11**, 560–566.
40 M. Amblard, J. A. Fehrentz, J. Martinez and G. Subra, *Mol. Biotechnol.*, 2006, **33**, 239–254.
41 S. Park, S. Gunasekera, T. L. Aboye and U. Goransson, *Int. J. Pept. Res. Ther.*, 2010, **16**, 167–176.
42 T. Takeuchi, M. A. Shuman and C. S. Craik, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11054–11061.
43 T. Chase and E. Shaw, *Biochemistry*, 1969, **8**, 2212–2224.
44 G. W. Jameson, D. V. Roberts, R. W. Adams, W. S. A. Kyle and D. T. Elmore, *Biochem. J.*, 1973, **131**, 107–117.
45 W. R. Greco and M. T. Hakala, *J. Biol. Chem.*, 1979, **254**, 2104–2109.
46 M. R. Krause, R. Goddard and S. Kubik, *J. Org. Chem.*, 2011, **76**, 7084–7095.
47 A. Brik, J. Alexandratos, Y. C. Lin, J. H. Elder, A. J. Olson, A. Wlodawer, D. S. Goodsell and C. H. Wong, *ChemBioChem*, 2005, **6**, 1167–1169.
48 A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623–1641.

2.2. Combinatorial Tuning of Peptidic Drug Candidates: High-Affinity Matriptase Inhibitors through Incremental Structure-Guided Optimization

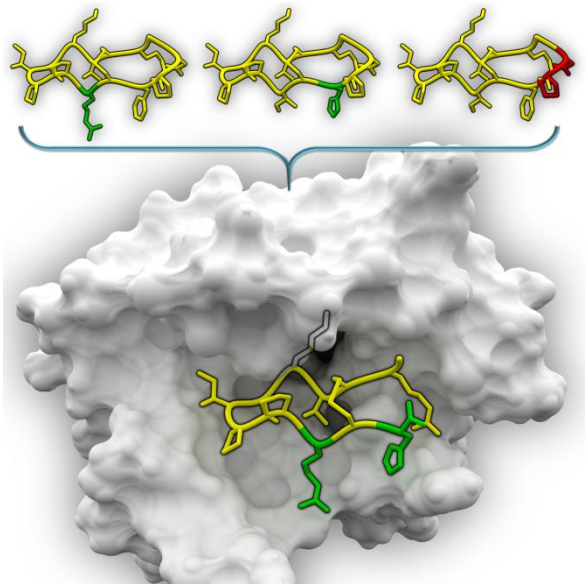
Authors:

Heiko Fittler, Olga Avrutina, Bernhard Glotzbach, Martin Empting, Harald Kolmar

Bibliographic Data:

Organic & Biomolecular Chemistry,
Volume 11, Issue 11, Page 1848-1857, 21 March 2013.
DOI: 10.1039/c3ob27469a
First published online: 29 January, 2013.

Graphical Abstract:



Contributions by H. Fittler

- Synthesized and isolated all peptidic compounds
- Performed inhibition assays for all compounds

Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Combinatorial tuning of peptidic drug candidates: high-affinity matriptase inhibitors through incremental structure-guided optimization†

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 1848

Heiko Fittler, Olga Avrutina, Bernhard Glotzbach, Martin Empting* and Harald Kolmar*

Received 19th September 2012,
Accepted 8th January 2013

DOI: 10.1039/c3ob27469a

www.rsc.org/obc

Herein we report a convenient strategy for the development of novel, highly-potent peptidic inhibitors of the trypsin-like serine protease matriptase based on the monocyclic variant of the sunflower trypsin inihibitor-1 (SFTI-1[*1,14*]). We screened SFTI-1[*1,14*] variants possessing incremental modifications of the parent peptide for beneficial binding properties. This compound library comprising 6 peptides and 16 triazole-containing peptidomimetics was established *via* structure-guided rational design and synthesized using a divergent strategy employing "copper-click" chemistry. The most favorable amino acid substitutions were combined in one framework yielding potent SFTI-1-derived matriptase inhibitor-1 (SDMI-1) and the truncated dodecapeptide variant (SDMI-2) with single-digit nanomolar inhibition constants. *In silico* studies indicated that the improved matriptase affinity compared to the parent peptide is caused by the successful establishment of additional favorable proton donor–acceptor interactions between basic inhibitor side chains and acidic residues on the surface of the target enzyme. SDMI-1 and 2 are potent inhibitors of the pharmaceutically relevant protease matriptase at a near physiological pH and, thus, may find applications in therapy or diagnostics.

Introduction

Trypsin is one of the most prominent digestive enzymes ubiquitously found in the small intestine of vertebrates.[1] Its intriguing molecular framework includes the famous triad Asp-His-Ser as a core feature implementing its proteolytic activity.[2] This prototypic architecture and the ability to cleave peptide bonds after basic residues constitute the structural and functional groundwork of a whole class of biocatalysts referred to as trypsin-like serine proteases.[3] Members of this enzyme family are involved in diverse biological processes and occur in soluble form or as membrane-anchored entities.[4] Type II transmembrane serine proteases (TTSP), for instance, are bound to the cell surface *via* the N-terminus and have been characterized as important mediators of the pericellular procession and activation of various effector molecules.[3–7] Active forms of peptide hormones, growth and differentiation factors, receptors, enzymes, and adhesion molecules are generated from inactive precursors through endoproteolytic cleavage by specific TTSPs.[4] Hence, they play crucial roles in the cellular development and maintenance of homeostasis.[3]

A well-studied example of a membrane-anchored trypsin-like serine protease with pharmaceutical relevance is matriptase.[8–18] It is widely expressed on the surface of epithelial cells in healthy tissue where its proteolytic activity is precisely regulated by natural protease inhibitors like the hepatocyte growth factor activator inhibitor-1 and 2 (HAI-1, HAI-2).[10,13] However, dysregulation of this physiological inhibitor–protease balance is believed to facilitate pathological processes. Indeed, a number of studies associate matriptase overexpression with the development and progression of epithelial tumors, as well as osteoarthritis and atherosclerosis.[14–18] Furthermore, Napp *et al.* observed pronounced *in vivo* matriptase activity in a murine orthotopic pancreatic tumor model and showed that the administration of active-site inhibitors significantly reduces proteolysis of the substrate analyte.[11] Hence, potent and selective matriptase inhibitors are of great therapeutic importance, and their development is a challenging task. To date, a number of small synthetic organic compounds as well as large antibody fragments exhibiting single-digit nanomolar to subnanomolar inhibition constants have been reported.[19–21] Very recently, Marsault

Clemens-Schöpf Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Petersenstr. 22, 64287 Darmstadt, Germany.
E-mail: M.Empting@Biochemie-TUD.de, Kolmar@Biochemie-TUD.de

†Electronic supplementary information (ESI) available: Tabulated data of ESI-MS analysis, RP-HPLC chromatograms, and plotted kinetic data of enzyme inhibition assays for all compounds. See DOI: 10.1039/c3ob27469a

and coworkers have demonstrated the applicability of peptidomimetics bearing a benzothiazole-based "serine trap" for the development of very potent matriptase inhibitors.[21] Reactive carbonyl functionality at the P_1 position allowed for the establishment of a transient covalent linkage between the ligand and the active-site serine.[21,25,26]

In this study, we present a convenient strategy for the development of non-covalent high-affinity matriptase inhibitors starting from a naturally occurring peptidic framework comprising a canonical protease inhibitor loop (Fig. 1). We used the rigid and well-defined tetradecapeptide framework of the sunflower trypsin inhibitor-1 (SFTI-1) as a starting point for structure-guided lead compound optimization. This molecular scaffold possesses a canonical (substrate-mimicking) loop for anti-proteolytic activity and has already been successfully used for the generation of inhibitors of chymotrypsin, elastase, cathepsin G, β-tryptase, proteinase K or kallikrein-related peptidase.[27–31] Moreover, unlike a majority of peptidic molecules, both bicyclic and monocyclic variants of SFTI-1 demonstrate a sufficient stability in blood serum (35–76 h).[30,32] Additionally, a latent inhibitory activity against matriptase in the nanomolar range (K_i = 1–150 nM) has been reported for the bicyclic peptide at the pH optimum of the TTSP (pH ~ 8.5–9).[22,33–36] Very recently, we showed that this moderate potency is essentially decreased in an environment with near physiological parameters (K_i = 1.1 μM at pH 7.6).[22] This is of particular importance for potential *in vivo* applications. A therapeutic/diagnostic agent must possess high activity under homeostasis or – in the case of tumor as well as inflammation-induced hypoxia – more acidic conditions.[37–41]

Using a structure-guided incremental optimization strategy we were able to generate SFTI-1 derivatives with single-digit

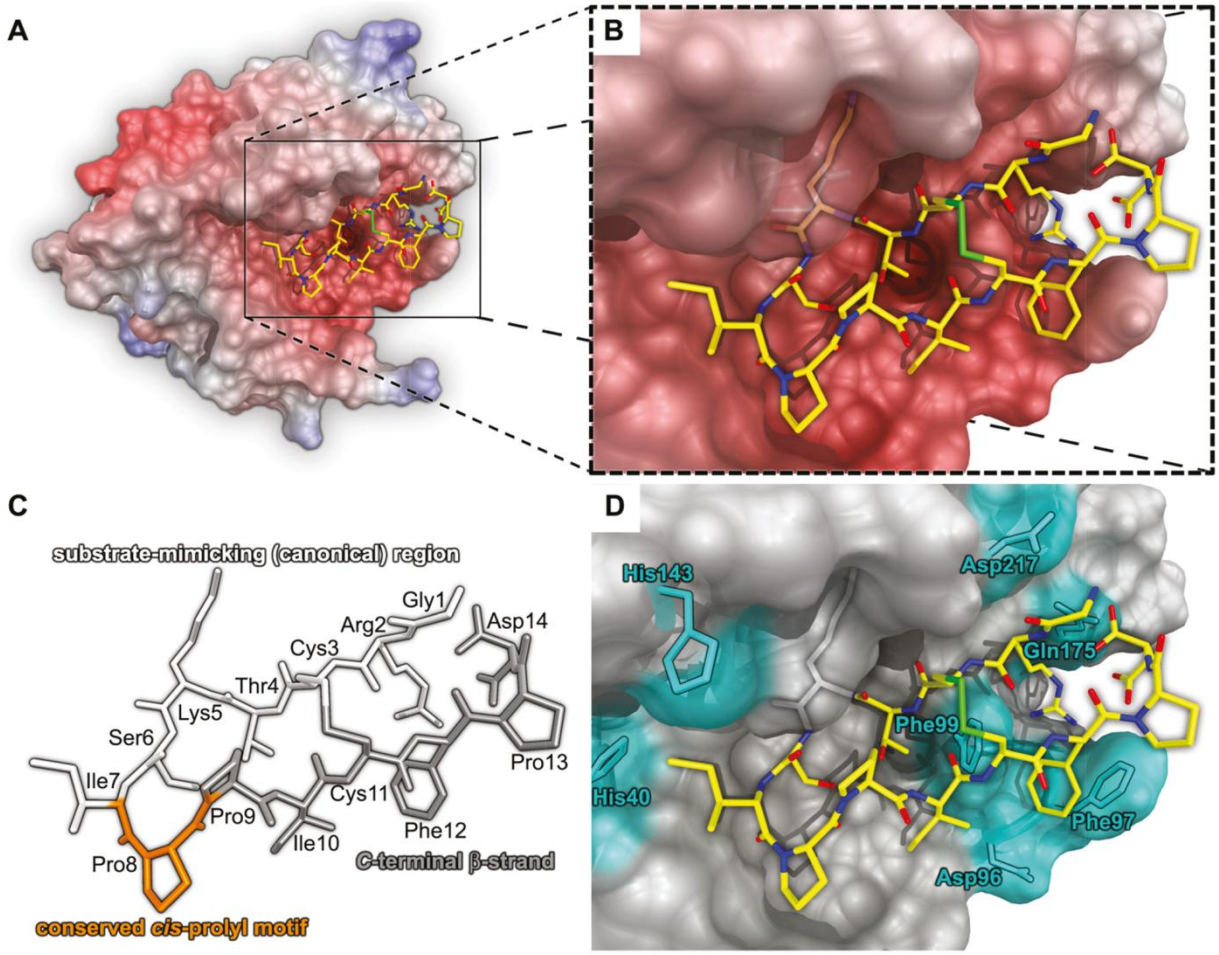


Fig. 1 Predicted structure of the SFTI-1[*1,14*] (**1**)–matriptase complex generated *via* molecular dynamics simulations followed by a two-step docking procedure.[22] Blue: nitrogen, green: sulfur, red: oxygen, yellow: carbon, hydrogen is omitted for clarity. Reported X-ray/NMR structures 3P8F and 1JBN were used as initial coordinates.[12,23] (A) Molecular surface of the catalytic domain of matriptase in complex with **1** (sticks). Color gradients (red to blue) indicate electrostatic surface potentials (negative to positive) at pH 7.4 calculated *via* the Particle Mesh Ewald method with a maximum value of 500 kJ mol^{-1}.[24] (B) Enlarged depiction of the active site. Regions concealing parts of the ligand are shown transparently. Bound inhibitor **1** is highlighted by a black outline. (C) Sequence and schematic representation of SFTI-1[*1,14*] (**1**). (D) Depiction of the active site indicating side chains of matriptase residues (cyan sticks, black outline) which have been considered within the presented approach for the structure-guided optimization of the SFTI-1 framework (sticks, black outline).

nanomolar K_i as well as an improved trypsin/matriptase selectivity profile.

Results and discussion

In our previous study, we reported that the monocyclic variant SFTI-1[*1,14*] possesses a slightly improved inhibition constant compared to the bicyclic wild type.[22] Thus, the open-chain derivative represents a promising starting point for the design of novel matriptase inhibitors. A combination of molecular dynamics simulation and docking experiments yielded an *in silico* SFTI-1[*1,14*]-matriptase complex, very similar to the reported wild-type inhibitor-protease X-ray structure (Fig. 1).[12,22]

Using these structural data and considering suggestions made by Yuan *et al.*,[12] we designed a minimal compound library **1**–**22** suitable for an initial search for incremental enhancements of the lead structure (step 1, Fig. 2). In a second step we investigated whether a combination of the individual beneficial modifications of the SFTI-1 framework results in an additive effect (step 2). Finally, we truncated the most promising variant to study the influence of the C-terminal residues on matriptase affinity and selectivity (step 3).

Step 1: tuning affinity of singular residues (increments)

Trypsin and matriptase share a very similar substrate spectrum. Thus, we rationalized that substitutions within the canonical region of SFTI-1[*1,14*] (Fig. 1C) would be detrimental towards improved affinity. However, we included the $P_{2'}$ position (Ile7) as a site for side-chain replacements within our molecular design (Fig. 2A). Histidines 40 and 143 of matriptase are in close proximity to this residue, therefore they might provide the possibility for favorable hydrogen bonding interactions with the ligand (Fig. 1D).

To reduce the synthetic expense and to cover an adequate structural and functional space, we set up a divergent synthetic procedure. Azide-bearing peptidic scaffolds **23** and **24** were assembled on the solid support using commercially available building blocks Fmoc-L-azidoalanine (Fmoc-Aza-OH) and Fmoc-L-azidohomoalanine (Fmoc-Aha-OH) (Fig. 2B). On-resin copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) with different alkyne components allowed for the facile installation of an amine, a carboxylic acid, the corresponding methyl ester and a phenyl functionality at position 7. This combinatorial approach is quite similar to the "tethered fragment" strategy previously described by Burke Jr. and coworkers.[42,43] Acidolytic cleavage from the solid support, oxidative macrocyclization and chromatographic isolation gave SFTI-1[*1,14*] derivatives **1**–**8** (Fig. 2C). Interestingly, both free carboxylic acids **3** and **4** as well as corresponding methyl esters **5** and **6** were selectively accessible from the respective open-chain methyl ester precursor peptides through the choice of macrocyclization conditions. The methyl carboxylate was hydrolyzed during disulfide bond formation at pH 8.4 giving free acids **3** and **4**. Setting the pH to 7.7, however, allowed for the preservation of the methyl ester group yielding the corresponding cystine-bridged products **5** and **6**.

Unfortunately, none of the position 7 variants **1**–**8** showed an improved matriptase affinity over SFTI-1[*1,14*] in enzyme inhibition assays (Table 1). Thus, we focused our further experiments on the optimization of positions 10 and 12.

In the ligand–receptor complex, the side chain of residue 10 of SFTI-1[*1,14*] is oriented towards a surface area of matriptase with a pronounced negative polarization (Fig. 1). The β-carboxylic group of the Asp96 contributes significantly to this environment. Yuan *et al.* suggested installing short basic side chain functionalities with low degrees of conformational freedom like aminoalanine or aminohomoalanine at position 10 of the inhibitor to establish favorable electrostatic interactions.[12] Nevertheless, initial molecular modeling implied that a linker of about four to five methylene units is needed to position a basic functionality in proximity to the β-carboxylic group of Asp96 in its native conformation present in the crystal structure (PDB-ID: 3P8F). Thus, we designed SFTI-1[*1,14*] derivatives **9**–**16** possessing planar and rigid triazolyl linkers of appropriate length to investigate the possibility for new favorable interactions between the ligand and the receptor (Fig. 2A).

Initially, synthetic routes for **9**–**16** were devised as modifications of the approach used for SFTI-1[*1,14*] variants **1**–**8**. While variants **10**, **12**, **14**, and **16** were readily accessible *via* this strategy (Fig. 2C), compounds **9**, **11**, **13**, and **15** could not be synthesized in this manner. An attempt to assemble the Aza10 analog of peptide-resin **25** on the solid support resulted in an undefined mixture of side products (data not shown). Nevertheless, we were able to synthesize compounds **9**, **11**, **12**, and **15** *via* the resin-bound pentapeptide intermediate **26** as depicted in Fig. 2D.

Interestingly, a significant increase of matriptase affinity over SFTI-1[*1,14*] was observed for four of the eight triazolyl-containing peptidomimetics **9**–**16** (Table 1).

The most pronounced enhancement within the set of inhibitors **9**, **10**, **13**, and **15** was detected for the aminomethyl-functionalized 1,2,3-triazole **9**. The difference of the free energies of binding/dissociation $\Delta_B G$ between compound **9** and SFTI-1[*1,14*] was calculated as 3.9 kJ mol^{-1}. Thus, additional favorable electrostatic inhibitor–enzyme interactions were successfully established by furnishing the peptide with a basic functionality at position 10. However, if length of the triazolyl linker is increased by only one methylene unit (**10**), the attractive contribution is significantly reduced. Furthermore, the installation of an acidic carboxy group at this position (**11**, **12**) is detrimental for matriptase inhibition. Noteworthy, a significant increase of inhibitor potency compared to SFTI-1[*1,14*] is also observed for the formally uncharged methyl ester derivative **13** although not as pronounced as in the case of the amino-functionalized variant **9**.

Additionally, we investigated the impact of basic canonical amino acids, lysine and arginine, at position 10 (**17**, **18**) on matriptase affinity. Surprisingly, compound **18** equipped with a flexible aliphatic linker demonstrated the highest potency of

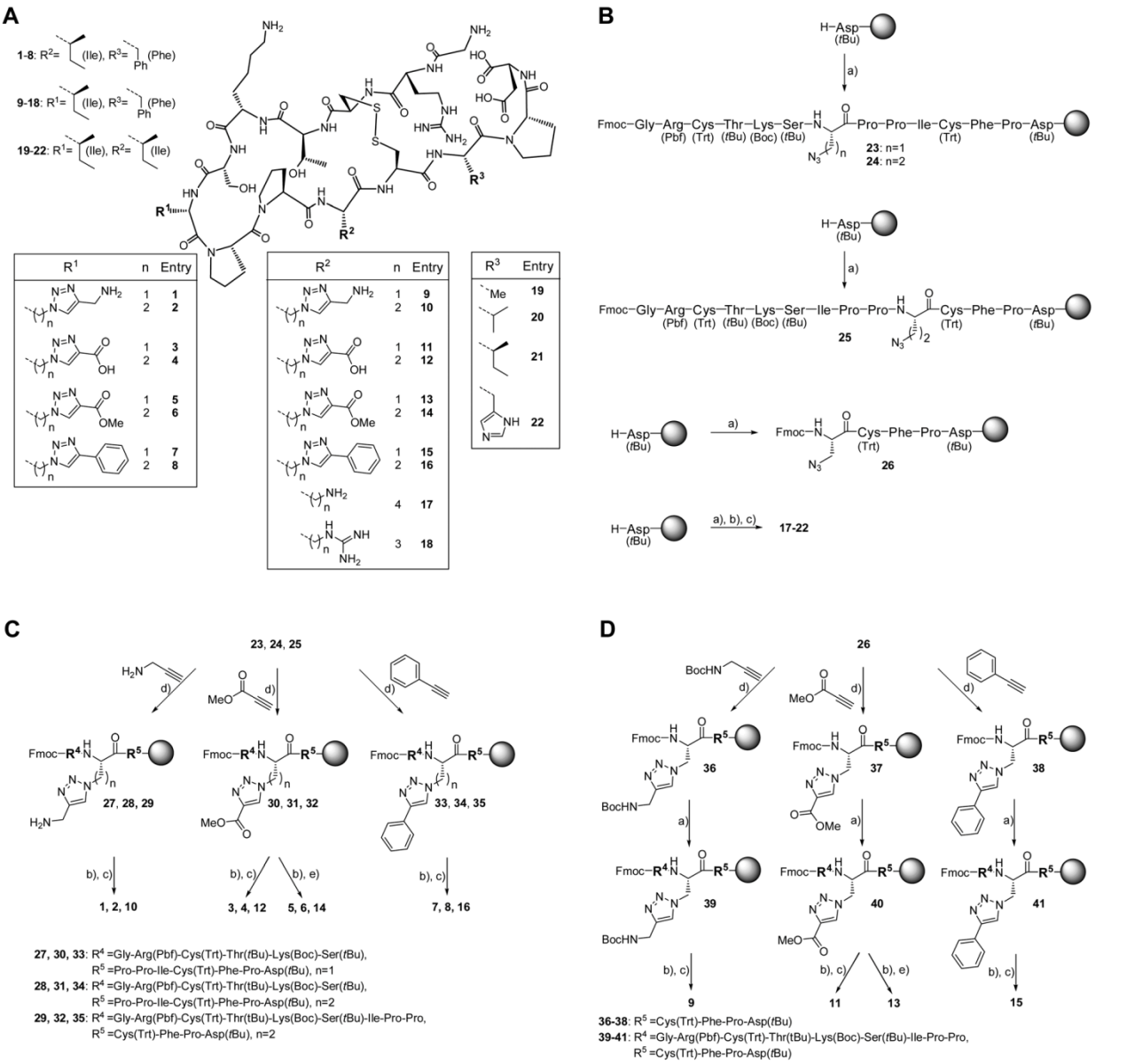


Fig. 2 Design and synthesis of a small compound library comprising SFTI-1[*1,14*] derivatives **1–22**. (A) Collection of triazolyl-containing peptides **1–16** as well as variants **17–22** with singular substitutions at positions 10 and 12 using canonical amino acids lysine, arginine, alanine, valine, isoleucine, or histidine. (B) Synthesis of precursors **23**, **24**, **25**, and **26** as well as inhibitors **17–22**. (C) and (D) Schematic depictions of synthetic routes yielding triazolyl-containing peptides **1–16**. Conditions: (a) microwave-assisted Fmoc-SPPS; (b) acidolytic cleavage from the solid support using TFA–H_2O–anisole–TES (47 : 1 : 1 : 1, v : v : v : v) and dithiothreitol (DTT), (c) air-mediated oxidative macrocyclization in 100 mM $(NH_4)_2CO_3$ aq. (pH = 8.4) at 1 mg peptide per mL; (d) on-resin CuAAC using given alkyne component (5 eq.), $CuSO_4{\cdot}5H_2O$ (20 mol%), sodium ascorbate (20 mol%), and *N*,*N*-diisopropylethylamine (DIEA, 8 eq.) in DMF at ambient temperature (overnight) (e) air-mediated oxidative macrocyclization in 100 mM $(NH_4)_2CO_3$ aq. (pH = 7.7) at 1 mg peptide per mL. All compounds **1–22** were isolated through preparative reversed-phase HPLC.

SFTI-1[*1,14*] variants **1–22** listed in Table 1. Due to the inherent degrees of conformational freedom of the respective amino acid side chains, a pronounced entropic penalty has been expected for compounds **17** and **18**.[12] Indeed, ε-amino-functionalized derivative **17** was slightly less potent than inhibitor **9** possessing a more rigid and restrained linker motif of comparable length. However, a double-digit nanomolar K_i corresponding to an increase of $\Delta_B G$ by 6.5 kJ mol^{-1} resulted from a singular amino acid substitution (**18**).

Finally, we investigated the impact of modifications at position 12 of the SFTI-1[*1,14*] framework on matriptase affinity. Yuan *et al.* described a pronounced undesirable entropic effect on Phe12 of SFTI-1 due to conformational restraints caused by Phe97 and Phe99 side chains forming the S_4 subsite of the

Table 1 Determined inhibition constants of compounds **1–22** against matriptase at pH 7.6 and calculated differences in free energies of binding/dissociation compared to SFTI-1[*1,14*]

Entry		K_i[a]/nM	Relative activity[b]	$\Delta_B G_{(x)} - \Delta_B G_{(SFTI-1[1,14])}$[c]/(kJ mol^{-1})
SFTI-1[*1,14*]		703 ± 87[d]	1	0
Pos. 7[e]	**1**	4892 ± 663	7.0	−4.8 ± 0.5
	2	3822 ± 494	5.4	−4.2 ± 0.4
	3	13 886 ± 1711	19.8	−7.4 ± 0.4
	4	2380 ± 291	3.4	−3.0 ± 0.4
	5	1629 ± 200	2.3	−2.1 ± 0.4
	6	10 857 ± 1349	15.4	−6.8 ± 0.4
	7	3342 ± 414	4.8	−3.9 ± 0.4
	8	1252 ± 155	1.8	−1.4 ± 0.4
Pos. 10[e]	**9**	148 ± 19	0.21	3.9 ± 0.4
	10	513 ± 66	0.73	0.8 ± 0.4
	11	46 287 ± 5728	65.8	−10.4 ± 0.4
	12	8096 ± 1011	11.5	−6.1 ± 0.4
	13	208 ± 27	0.30	3.0 ± 0.4
	14	2224 ± 277	3.2	−2.9 ± 0.4
	15	556 ± 70	0.79	0.6 ± 0.4
	16	4750 ± 585	6.8	−4.7 ± 0.4
	17	232 ± 29	0.33	2.7 ± 0.4
	18	50.6 ± 6.5	0.072	6.5 ± 0.4
Pos. 12[e]	**19**	1614 ± 206	2.3	−2.1 ± 0.4
	20	580 ± 72	0.83	0.5 ± 0.4
	21	319 ± 40	0.45	2.0 ± 0.4
	22	206 ± 27	0.29	3.0 ± 0.5

[a] Determined as described in the Experimental section. [b] Relative activity given as the ratio K_i of the corresponding compound/K_i of SFTI-1[*1,14*]. [c] $\Delta_B G_{(x)}$ refers to the free energy of binding/dissociation of compound x. $\Delta_B G$ were calculated from respective K_i using $\Delta_B G = -RT\ln K_i$.[22,44] Errors of the given differences $\Delta_B G_{(x)} - \Delta_B G_{(SFTI-1[1,14])}$ were calculated by propagation of errors. [d] As published before.[22] [e] The modified position of the SFTI-1[*1,14*] framework is given.

enzyme (Fig. 1D).[12] As a consequence, substitution of Phe12 by non-natural amino acids possessing larger aromatic side chains has a detrimental effect on bioactivity.[34] Thus, we decided to replace the native phenylalanine residue by smaller aliphatic, hydrophobic, as well as heteroaromatic amino acids alanine, valine, isoleucine, and histidine (Fig. 2A). Respective compounds **19–22** were synthesized through routine microwave-assisted Fmoc-SPPS followed by oxidative macrocyclization and chromatographic isolation (see ESI†).

Indeed, we observed a beneficial effect of Phe12Ile (**21**) as well as Phe12His (**22**) substitutions on matriptase inhibition in our *in vitro* assays. In particular, the imidazole functionality at position 12 (**22**) facilitated an improved interaction between the SFTI-1 framework and the surface of the enzyme of about 3 kJ mol^{-1}.

Step 2: combination of beneficial increment modifications

Encouraged by the enhanced matriptase affinities observed for compounds **9**, **18** and **22**, we combined the singular modifications of position 10 (**9**, **18**) with the described Phe12His substitution (**22**). The resulting constructs **42** and **43** (Fig. 3) were synthesized either *via* microwave-assisted Fmoc-SPPS (**42**) or by using a strategy analogous to that described for compounds **9–16** (**42**) (compare Fig. 2D).

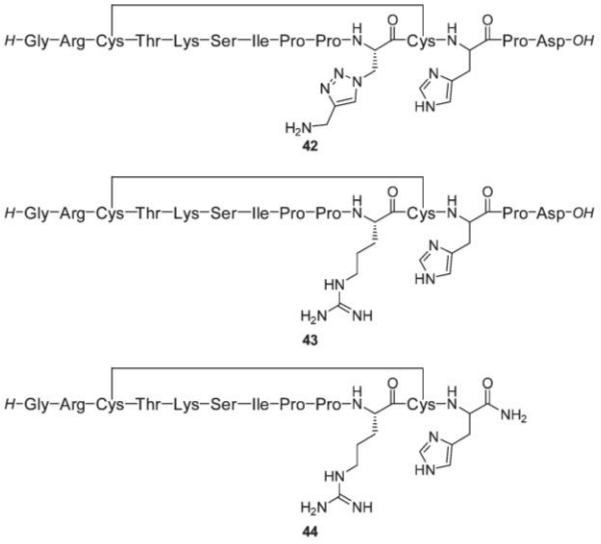


Fig. 3 Structures of triazolyl-containing SFTI-1[*1,14*] derivative **42**, SFTI-1 derived matriptase inhibitor-1 (SDMI-1) **43** and SFTI-1 derived matriptase inhibitor-2 (SDMI-2) **44**.

Inhibitors **42** and **43** demonstrated a further increased matriptase affinity compared to SFTI-1[*1,14*] derivatives **9**, **18**, and **22** (Table 2). Our calculations suggest that the observed improvement in $\Delta_B G$ relies predominantly on an additive effect. The determined individual favorable contributions of compounds **18** and **22** sum up to 9.6 ± 0.6 kJ mol^{-1}, whereas, the difference of free energies of binding/dissociation between **43** and SFTI-1[*1,14*] was calculated as 10.3 ± 0.4 kJ mol^{-1}. The difference between these two values could be accounted for a synergistic behavior.[44] However, the observed effect is negligible considering the respective estimated uncertainties. In the case of inhibitor **42** even an unfavorable non-additive portion was detected for the installed modifications as the calculated increase of $\Delta_B G$ was smaller than the sum of the incremental improvements (Table 2).

Nevertheless, tetradecapeptide **43** which we refer to as the SFTI-1-derived matriptase inhibitor-1 (SDMI-1) possesses a low inhibition constant of 11 nM near the physiological pH. Furthermore, in the enzyme assay conducted near the pH optimum of matriptase (pH = 8.5) a single-digit nanomolar K_i (1.1 nM) was observed.

Step 3: truncation of most promising variant

In a very recent study we performed a structural analysis of several monocyclic triazole-bridged SFTI-1 derivatives *in silico*.[22] The yielded results indicated a pronounced influence of the terminal regions on matriptase affinity. An entropic penalty caused by the fixation of the Phe12 residue of the parent bicyclic peptide through aromatic side chains of the enzyme has already been described and addressed in the presented work.[12] However, additional conformational constraints of residues Pro13 and Asp14 forming the secondary loop may cause unfavourable entropic contributions upon binding to

Table 2 Determined inhibition constants of compounds **42–44** against matriptase at pH 7.6 (8.5), calculated differences in free energies of binding/dissociation compared to SFTI-1[*1,14*] as well as trypsin affinity and selectivity at pH 7.6[a]

Entry	Matriptase K_i[b]/nM	Relative activity	$\Delta_B G_{(x)} - \Delta_B G_{(SFTI-1[1,14])}$/ (kJ mol^{-1})	Sum of increment contributions[c]/kJ mol^{-1}	Trypsin K_i/nM	Selectivity[d]
SFTI-1[*1,14*]	703 ± 87[e]	1	0	n.a.	0.21 ± 0.03[e]	0.0003
42	89.5 ± 11.1	0.13	5.1 ± 0.4	6.9 ± 0.6	9.2 ± 2.1	0.1
43 (SDMI-1)	11.0 ± 1.4 (1.1 ± 0.3)	0.016	10.3 ± 0.4	9.6 ± 0.6	11.2 ± 2.0	1.0
44 (SDMI-2)	6.2 ± 1.2 (2.3 ± 0.8)	0.009	11.7 ± 0.6	n.a.	38.0 ± 11.1	6.1

[a] Experiments were performed as described in the footnote of Table 1 and in the Experimental section. [b] Values in brackets correspond to K_i determined at pH 8.5. [c] Errors were calculated by propagation of errors. [d] Calculated as the ratio of K_i against matriptase (pH 7.6)/K_i against trypsin (pH 7.6). [e] As published before.[22]

matriptase. According to our calculations interaction with trypsin is not affected in this manner. As a consequence, the truncation of the two C-terminal residues might have a beneficial effect on matriptase inhibition. Additionally, we rationalized that removing the negative charge of the terminal carboxy group by introduction of a C-terminal amide might enhance the overall charge complementarity between the target enzyme and the inhibitor.

Inspired by these considerations, we synthesized the monocyclic dodecapeptide **44** as described in the Experimental section (Fig. 3). Indeed, we observed an additional minor improvement of inhibitor potency resulting in a K_i of 6.2 ± 1.2 nM. This SFTI-1-derived matriptase inhibitor-2 (SDMI-2) exhibits a 2.3 ± 0.8 nM inhibition constant at pH 8.5 and shows an improved selectivity towards matriptase over trypsin (6-fold). To our knowledge, **43** and **44** are to date the most potent Bowmann–Birk matriptase inhibitors described.

Finally, we modeled the inhibitor–protease complex for SDMI-2 (**44**) based on the *in silico* coordinates of SFTI-1[*1,14*] and matriptase described above (see Fig. 1).[22] First, the amino acid exchanges Ile10Arg and Phe12His were introduced into the parent model. Then, residues Pro13 and Asp14 were replaced by a C-terminal amide. Optimization of side-chain geometry and subsequent energy minimization yielded a structure quite similar to the corresponding SFTI-1 co-crystal (3P8F, Fig. 4A). Nevertheless, additional favorable interactions between **44** and matriptase were observed (Fig. 4B). The side chain of Asp96 was reoriented to form a bidentate hydrogen bond with the guanidinyl functionality of Arg10. Furthermore, the possibility for a proton donor–acceptor interaction between the backbone carbonyl oxygen of Asp96 and the ε^2-nitrogen of His12 was detected. However, an X-ray structure of the SDMI-2–matriptase complex is needed to undoubtedly prove whether the predicted *in silico* coordinates are valid. Yuan *et al.* suggested to use short basic residues like aminohomoalanine at position 10 to establish an effective electrostatic interaction with Asp96 while attenuating entropic penalties arising from extended flexible linkers.[12] However, it remains to be tested whether such a short side chain is sufficient to position a basic functionality in proximity to the carboxylic group of Asp96.

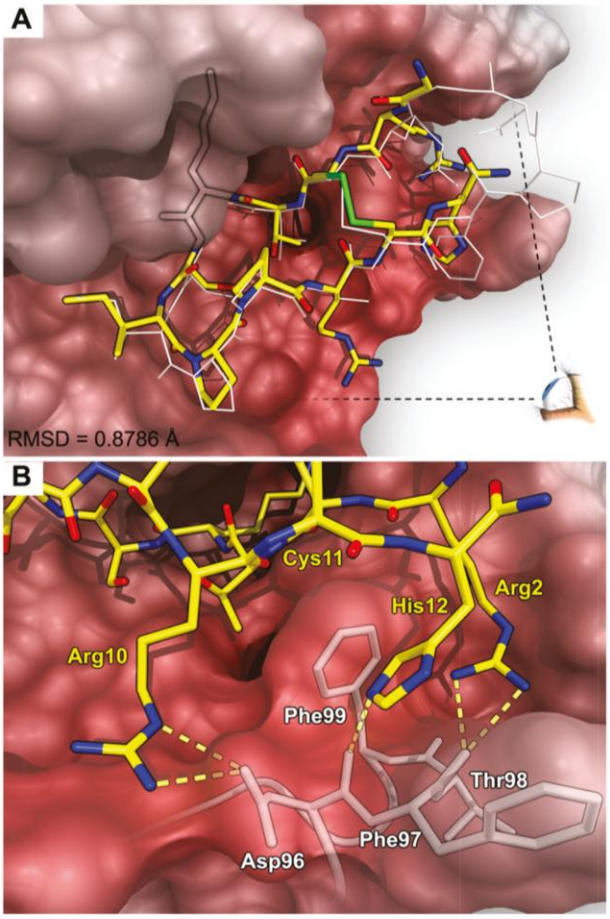


Fig. 4 (A) Predicted structure of SFTI-1 derived matriptase inhibitor-2 (SDMI-2) **44** (thick colored sticks) upon binding to matriptase (surface). The SFTI-1-matriptase X-ray structure (3P8F) is shown as thin white sticks and the calculated peptide backbone root mean square deviation (RMSD) is given for both complexes. The angle of view for the picture below is indicated. (B) Close-up of the predicted inhibitor **44**–enzyme complex. Important matriptase residues are depicted as transparent white sticks with black outlines and putative hydrogen bond interactions are shown as dashed yellow lines. Blue: nitrogen, green: sulfur, red: oxygen, yellow: carbon, hydrogen is omitted for clarity. Color gradients (red to blue) indicate electrostatic surface potentials (negative to positive) at pH 7.6 calculated *via* the Particle Mesh Ewald method with a maximum value of 500 kJ mol^{-1}.[24]

Conclusion

In this study we present a convenient strategy for the development of novel, highly potent inhibitors of matriptase based on the SFTI-1 framework. Using a divergent synthetic route we screened 22 compounds for beneficial modifications of the parent peptide. Combination of the most favorable amino acid substitutions and subsequent C-terminal truncation yielded SFTI-1 derived matriptase inhibitors-1 and 2 (SDMI-1 and 2) with inhibition constants in the low nanomolar to single-digit nanomolar range. The structure-guided design of the initial peptide/peptidomimetic library as well as the shortening of the amino acid sequence were inspired by *in silico* experiments which were used as an idea generator towards beneficial modifications of the wild-type compound.[12,22] Thus, an inhibitor possessing only twelve residues as well as an inverted trypsin–matriptase selectivity in favor of the latter enzyme was developed.

Although the applied synthetic strategy using "copper-click" chemistry was applicable for the assembly of the majority of the triazolyl-containing peptidomimetics **1**–**16**, it was incompatible with the use of non-natural amino acid azidoalanine at position 10 (compounds **9**, **11**, **13**, and **15**). In our previous works, we did not observe significant limitations using Fmoc-Aza-OH as a building block for microwave assisted Fmoc-SPPS.[45] However, the unexpected restricted applicability of the unprotected azide functionality under the described conditions was circumvented through a modified synthetic route.

Novel compounds **43** and **44** are potent inhibitors of matriptase in a pH range near the physiological one and, thus, may find applications in therapy or diagnostics. Notably, the activity of these SFTI-1 derivatives composed exclusively of natural amino acids is not in the range of recently reported peptidomimetics bearing reactive carbonyl functionalities.[21] Since only two residues have been furnished with new side-chain functionalities, the SFTI-1 framework provides further potential for the establishment of unique attractive interactions with matriptase and, therefore, improvement of affinity and selectivity. Additionally, our ongoing studies show that **43** and **44** can serve as versatile platforms for conjugation with reporter molecules without sacrificing bioactivity. This strategy would allow for *in vivo* imaging experiments using for example PET/SPECT or NIRF methodologies.[46–49]

Nevertheless, further selectivity tests using an extended set of relevant trypsin-like proteases as well as characterization of these two peptidic inhibitors *in vivo* are issues that have to be addressed in the future.

Experimental

General

All chemicals and solvents were purchased from Bachem, Iris Biotech, Novabiochem, Sigma-Aldrich, Rapp Polymere, Roth or Varian (Agilent).

Analytical and semi-preparative RP-HPLC were performed on a Varian 920-LC system and a Varian 940-LC system, respectively. A Phenomenex Hypersil 5u BDS C18 LC column (150 × 4.6 mm, 5 μm, 130 Å) and a Phenomenex Luna 5u C18 LC column (250 × 12.20 mm, 5 μm, 100 Å) were used as stationary phases. The eluent system consisted of eluent A (0.1% aq. TFA) and eluent B (90% aq. acetonitrile containing 0.1% TFA).

ESI mass spectra were recorded with a Shimadzu LCMS-2020 using a Phenomenex Jupiter 5u C4 LC column (50 × 1 mm, 5 μm, 300 Å) as well as a binary eluent system consisting of eluent A (0.1% aq. formic acid, LC-MS grade) and eluent B (100% acetonitrile containing 0.1% formic acid, LC-MS grade).

NMR spectra were recorded in DMSO-d_6 on a Bruker DRX 500 instrument (500 MHz) using a Shigemi 5 mm symmetrical NMR microtube (magnetic susceptibility matched to DMSO).

Peptides were synthesized using a Liberty 12-channel automated peptide synthesizer on a Discover SPS microwave peptide synthesizer platform (CEM) following the Fmoc strategy.

Chemistry

General procedures of peptide synthesis. All peptides bearing a C-terminal carboxy acid were assembled on a TentaGel S AC resin (Rapp Polymere) or a 2-chlorotrityl chloride resin (Iris Biotech) preloaded with Fmoc-L-Asp(*t*Bu). Corresponding peptide amides were assembled on an AmphiSpheres 40 RAM resin (Agilent).

Loading of 2-chlorotrityl chloride resin was performed manually as follows. A solution of 103 mg Fmoc-L-Asp(*t*Bu)-OH (0.25 mmol) and 34 μL *N*,*N*-diisopropylethylamine (DIEA, 1 mmol) in a minimal amount of dichloromethane (DCM) was added to the resin. The resulting mixture was shaken for 2 h at ambient temperature. The solution was removed by filtration and the loaded resin was washed with DCM–methanol–DIEA (17 : 2 : 1; 3×), DCM (3×), dimethylformamide (DMF, 3×) and DCM (3×).

Canonical amino acids were attached by double or triple coupling using 4 equiv. of the corresponding amino acid, 3.9 equiv. of 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronoium hexafluorophosphate (HBTU) and 8 equiv. of DIEA, or, in the case of cysteine, 3–4 equiv. of 2,4,6-trimethylpyridine (collidine). Arginine and cysteine were coupled using a two-step microwave program: 1. RT, 0 W, 25 min; 2. 75 °C, 25 W, 0.5 min (Arg) and 1. RT, 0 W, 2 min; 2. 50 °C, 25 W, 4 min (Cys), respectively. All other amino acids were coupled using a standard microwave program: 75 °C, 21 W, 5 min.

Non-natural azide-bearing building blocks Fmoc-L-Aza-OH and Fmoc-L-Aha-OH were attached *via* double coupling using 2 equiv. of the corresponding amino acid, 1.9 equiv. HATU, and 4 equiv. DIEA and a two-step microwave program (1. 60 °C, 30 W, 45 min, 2. 75 °C, 20 W, 5 min).

Fmoc deprotection was achieved in two steps by reaction with 20% piperidine in DMF at 75 °C, 42 W for 0.5 min (initial

deprotection) followed by a second deprotection step with 20% piperidine in DMF at 75 °C, 42 W for 3 min.

Cleavage of peptides from the solid support and removal of side-chain protecting groups were achieved *via* acidolysis using a standard cleavage cocktail consisting of trifluoroacetic acid (TFA)–H_2O–anisole–triethylsilane (TES) (47 : 1 : 1 : 1, v : v : v : v) and DTT to suppress unwanted oxidation. The resulting reaction mixture was shaken for 3 h at RT followed by precipitation and subsequent washing (3×) with methyl tertiary butyl ether (MTBE) to yield crude linear peptides.

Air-mediated oxidative macrocyclization of the crude peptides was conducted in 100 mM $(NH_4)_2CO_3$ aq. (1 mg mL^{-1}, pH = 8.4) and monitored by analytical RP-HPLC. After complete conversion (1–7 days) the solvent was removed by freeze-drying to yield the crude peptides. To suppress unwanted saponification of the methyl ester of compounds **5**, **6**, **13**, and **14**, a 100 mM $(NH_4)_2CO_3$ aq. buffer (pH 7.7) was used (Fig. 2).

Triazolyl-containing peptides 1–8, 10, 12, 14, and 16. Azide-bearing peptide resins **23**, **24**, and **25** were assembled on a TentaGel S AC-Asp(*t*-Bu) Fmoc resin (loading: 0.21 mmol g^{-1}) according to the described procedures on a 0.25 mmol scale, washed (3× with DCM and 3× with ether) and dried in an exsiccator.

On-support copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) was conducted with 0.05 mmol of intermediates **23**, **24**, and **25** and the respective alkyne component (propargylamine, methyl propiolate, or phenylacetylene) according to the scheme laid out in Fig. 2C. A solution of 5 equiv. alkyne, 1 equiv. copper(II) sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$), 1 equiv. sodium ascorbate (NaAsc) and 8 equiv. DIEA in 5 mL argon-flushed DMF was added to the peptide resin and shaken at ambient temperature overnight. Then, the solution was removed by filtration and the peptide resin was washed with methanol (3×), 0.5% sodium diethyldithiocarbamate in DMF (w/v, 3×), DMF (3×) and DCM (3×) yielding intermediates **27–35**.

After Fmoc deprotection, acidolytic cleavage from the solid support, ether precipitation, and oxidative disulfide formation, macrocyclic peptides were isolated *via* semi-preparative RP-HPLC. This gave compounds **1–8**, **10**, **12**, **14**, and **16** as white solids in the following yields. **1**: 13.5 mg (17%); **2**: 14.9 mg (18.6%); **3**: 9.3 mg (11.6%); **4**: 6.3 mg (7.8%); **5**: 2.4 mg (3%); **6**: 2 mg (2.5%); 7: 0.5 mg (0.6%); **8**: 5.1 mg (6.2%); **10**: 4.6 mg (5.8%); **12**: 4 mg (5%); **14**: 4 mg (4.9%); **16**: 2 mg (2.4%). For analytical data, please refer to the ESI.†

Triazolyl-containing peptides 9, 11, 13, and 15. Azide-bearing peptide resin **26** was assembled on a TentaGel S AC-Asp(*t*-Bu) Fmoc resin (loading: 0.21 mmol g^{-1}) according to the described procedures on a 0.25 mmol scale, washed (3× with DCM and 3× with ether) and dried in an exsiccator.

On-support copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) was conducted with 0.05 mmol of intermediates **26** and the respective alkyne component (*N*-Boc-propargylamine, methyl propiolate, or phenylacetylene) according to the scheme laid out in Fig. 2D. A solution of 5 eq. alkyne, 1 eq. copper(II) sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$), 1 eq. sodium ascorbate (NaAsc) and 8 eq. DIEA in 5 mL argon-flushed DMF was added to the peptide resin and shaken at ambient temperature overnight. Then, the solution was removed by filtration and the peptide resin was washed with methanol (3×), 0.5% sodium diethyldithiocarbamate in DMF (w/v, 3×), DMF (3×) and DCM (3×) yielding intermediates **36–38**.

The Fmoc protecting group was removed according to the described procedure. The remaining N-terminal residues were attached to each triazolyl-containing peptide resin yielding intermediates **39–41**. After Fmoc deprotection, acidolytic cleavage from the solid support, ether precipitation, and oxidative disulfide formation, macrocyclic peptides were isolated *via* semi-preparative RP-HPLC. This gave compounds **9**, **11**, **13**, and **15** as white solids in the following yields. **9**: 16.8 mg (21.2%); **11**: 6.6 mg (8.2%); **13**: 4.1 mg (5.1%); **15**: 4.2 mg (5.1%). For analytical data, please refer to the ESI.†

Triazolyl-containing peptide 42. Inhibitor **42** was synthesized following a strategy similar to that described for compounds **9**, **11**, **13**, and **15**. First, peptide-resin intermediate Fmoc-Aza-Cys(Trt)-His(Trt)-Pro-Asp(*t*Bu)-resin was assembled on a 2-chlorotrityl chloride resin preloaded with Fmoc-Asp (*t*Bu) (loading: 0.23 mmol g^{-1}) on a 0.05 mmol scale. Then, a solution of 39 mg *N*-Boc-propargylamine (0.25 mmol), 14 mg $CuSO_4 \cdot 5H_2O$ (0.05 mmol), 10 mg NaAsc (0.05 mmol) and 70 μL DIEA (0.4 mmol) in 5 mL argon-flushed DMF was added to the peptide resin and shaken at ambient temperature overnight. The solution was removed by filtration and the peptide resin was washed with methanol (3×), 0.5% sodium diethyldithiocarbamate in DMF (w/v, 3×), DMF (3×) and DCM (3×).

The Fmoc protecting group was removed according to the described procedure. The remaining N-terminal residues were attached to the triazolyl-containing peptide resin. After Fmoc deprotection, acidolytic cleavage from the solid support, ether precipitation, and oxidative disulfide formation, macrocyclic peptide **42** was isolated *via* semi-preparative RP-HPLC. This gave 0.9 mg (1.2%) of pure **42** as a white solid. For analytical data, please refer to the ESI.†

Peptides 17–22, and 43. Linear precursors were synthesized on a 2-chlorotrityl chloride resin preloaded with Fmoc-Asp-(*t*Bu) (loading: 0.23 mmol g^{-1}) on a 0.05 mmol scale.

After acidolytic cleavage from the solid support, ether precipitation, and oxidative disulfide formation, macrocyclic peptides were isolated *via* semi-preparative RP-HPLC. This gave compounds **17–22** and **43** as white solids in the following yields. **17**: 3 mg (3.9%); **18**: 6.2 mg (7.9%); **19**: 17.4 mg (23.9%); **20**: 11.2 mg (15.1%); **21**: 23.2 mg (31%); **22**: 9.2 mg (12.1%); **43**: 0.9 mg (1.2%). For analytical data, please refer to the ESI.†

Inhibitor 44. The linear precursor of **44** was synthesized on an AmphiSpheres 40 RAM resin (0.44 mmol g^{-1}) on a 0.1 mmol scale. After acidolytic cleavage from the solid support, ether precipitation, and oxidative disulfide formation, the macrocyclic peptide was isolated *via* semi-preparative RP-HPLC. This gave 10.5 mg (7.8%) of **44** as a white solid. For analytical data, please refer to the ESI.†

In vitro assays

Enzymes. Recombinant production, autocatalytic activation and purification of matriptase were conducted as previously described.[22] Bovine trypsin was purchased from Sigma-Aldrich. Active-site titration and determination of Michaelis–Menten constants K_M for the chromogenic substrate Boc-QAR-*p*NA were performed as reported earlier.[22]

Data collection and analysis. All experiments were performed in triplicate. Active-site titrated matriptase (final concentration: [E] = 0.9 nM) or trypsin (final concentration: [E] = 0.5 nM) were incubated with the respective inhibitor in different concentrations [I] at pH 7.6 or 8.5 for 30 min. Then, the chromogenic substrate Boc-QAR-*p*NA (final concentration: [S] = 250 μM) was added. The residual proteolytic activity (v/v_0) was determined by monitoring the absorption of the corresponding samples in 96-well plates (NUNC, round bottom, clear) at 405 nm in 60 s intervals over 30 min at RT using the Tecan GENios microplate reader. Slopes of the initial reaction velocities (steady-state) were calculated and normalized to the data of the uninhibited enzymatic hydrolysis of Boc-QAR-*p*NA. The determined v/v_0 values were plotted against the concentration of the inhibitor. The resulting dose–response curves were fitted either by eqn (1) or in the case of tight-binding inhibition (trypsin: **42**; matriptase: **43** at pH 8.5, **44** at pH 7.6 and 8.5) by eqn (2) through non-linear regression.

$$\frac{v}{v_0} = \left(1 + \frac{[\mathrm{I}]}{\mathrm{IC}_{50}}\right)^{-1} \quad (1)$$

$$\frac{v}{v_0} = 1 - \frac{\left([\mathrm{E}] + [\mathrm{I}] + K_\mathrm{i}^\mathrm{app}\right) - \sqrt{\left([\mathrm{E}] + [\mathrm{I}] + K_\mathrm{i}^\mathrm{app}\right)^2 - 4[\mathrm{E}][\mathrm{I}]}}{2[\mathrm{E}]} \quad (2)$$

Substrate-independent inhibition constants K_i were calculated from respective IC_{50} values or apparent (substrate-dependent) inhibition constants $K_\mathrm{i}^\mathrm{app}$ using eqn (3) and (4), respectively.

$$K_\mathrm{i} = \mathrm{IC}_{50}\left(1 + \frac{[\mathrm{S}]}{K_\mathrm{M}}\right)^{-1} \quad (3)$$

$$K_\mathrm{i} = K_\mathrm{i}^\mathrm{app}\left(1 + \frac{[\mathrm{S}]}{K_\mathrm{M}}\right)^{-1} \quad (4)$$

Errors were calculated through propagation of errors (compare Tischler *et al.*).[50]

In silico procedures

General. All *in silico* experiments were performed on an Intel® Core™ i7–2600 workstation with 8 virtual cores using the YASARA structure package (YASARA Biosciences) and the YASARA2 force field.[51] Graphical content was generated with YASARA and POVRay.[51,52]

SDMI-2–matriptase complex. Prediction of an SFTI-1[*1,14*]–matriptase complex based on the X-ray structure of the SFTI-1–matriptase complex (PDB ID: 3P8F) has been recently described.[12,22] Ile10 and Phe12 were replaced by Arg10 and His12 within the calculated structure of the open-chain variant. Residues Pro13 and Asp14 were deleted and the C-terminal amide functionality was installed. Then, Arg10 and His12 side-chain geometries were optimized using a rotamer library and steepest decent minimization.[53,54] A global energy minimization with an increased scaling factor for non-bonded interactions (1.2 instead of 1.0) was performed. Finally, a second energy minimization was performed with a reset scaling factor for non-bonded interactions (1.0) to yield the SDMI-2–matriptase complex depicted in Fig. 4.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft through grant Ko 1390/9-1, by BMBF, and LOEWE-Soft Control.

References

1 W. R. Rypniewski, A. Perrakis, C. E. Vorgias and K. S. Wilson, *Protein Eng.*, 1994, 7, 57–64.
2 A. Schmidt, C. Jelsch, P. Ostergaard, W. Rypniewski and V. S. Lamzin, *J. Biol. Chem.*, 2003, **278**, 43357–43362.
3 T. M. Antalis, T. H. Bugge and Q. Wu, *Prog. Mol. Biol. Transl. Sci.*, 2011, **99**, 1–50.
4 T. M. Antalis, M. S. Buzza, K. M. Hodge, J. D. Hooper and S. Netzel-Arnett, *Biochem. J.*, 2010, **428**, 325–346.
5 T. H. Bugge, T. M. Antalis and Q. Wu, *J. Biol. Chem.*, 2009, **284**, 23177–23181.
6 J. D. Hooper, J. A. Clements, J. P. Quigley and T. M. Antalis, *J. Biol. Chem.*, 2001, **276**, 857–860.
7 S. Netzel-Arnett, J. D. Hooper, R. Szabo, E. L. Madison, J. P. Quigley, T. H. Bugge and T. M. Antalis, *Cancer Metastasis Rev.*, 2003, **22**, 237–258.
8 M. S. Buzza, S. Netzel-Arnett, T. Shea-Donohue, A. P. Zhao, C. Y. Lin, K. List, R. Szabo, A. Fasano, T. H. Bugge and T. M. Antalis, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 4200–4205.
9 K. List, C. C. Haudenschild, R. Szabo, W. J. Chen, S. M. Wahl, W. Swaim, L. H. Engelholm, N. Behrendt and T. H. Bugge, *Oncogene*, 2002, **21**, 3765–3779.
10 R. Szabo, J. P. Hobson, K. List, A. Molinolo, C. Y. Lin and T. H. Bugge, *J. Biol. Chem.*, 2008, **283**, 29495–29504.
11 J. Napp, C. Dullin, F. Muller, K. Uhland, J. B. Petri, A. van de Locht, T. Steinmetzer and F. Alves, *Int. J. Cancer*, 2010, **127**, 1958–1974.
12 C. Yuan, L. Q. Chen, E. J. Meehan, N. Daly, D. J. Craik, M. D. Huang and J. C. Ngo, *BMC Struct. Biol.*, 2011, 11.
13 M. D. Oberst, L. Y. Chen, K. Kiyomiya, C. A. Williams, M. S. Lee, M. D. Johnson, R. B. Dickson and C. Y. Lin, *Am. J. Physiol. Cell Physiol.*, 2005, **289**, C462–C470.
14 K. List, R. Szabo, A. Molinolo, V. Sriuranpong, V. Redeye, T. Murdock, B. Burke, B. S. Nielsen, J. S. Gutkind and T. H. Bugge, *Genes Dev.*, 2005, **19**, 1934–1950.
15 K. Uhland, *Cell. Mol. Life Sci.*, 2006, **63**, 2968–2978.

16 K. List, *Future Oncol.*, 2009, **5**, 97–104.
17 J. M. Milner, A. Patel, R. K. Davidson, T. E. Swingler, A. Desilets, D. A. Young, E. B. Kelso, S. T. Donell, T. E. Cawston, I. M. Clark, W. R. Ferrell, R. Plevin, J. C. Lockhart, R. Leduc and A. D. Rowan, *Arthritis Rheum.*, 2010, **62**, 1955–1966.
18 I. Seitz, S. Hess, H. Schulz, R. Eckl, G. Busch, H. P. Montens, R. Brandl, S. Seidl, A. Schomig and I. Ott, *Arterioscl. Thromb. Vasc. Biol.*, 2007, **27**, 769–775.
19 T. Steinmetzer, A. Schweinitz, A. Sturzebecher, D. Donnecke, K. Uhland, O. Schuster, P. Steinmetzer, F. Muller, R. Friedrich, M. E. Than, W. Bode and J. Sturzbecher, *J. Med. Chem.*, 2006, **49**, 4116–4126.
20 C. J. Farady, J. Sun, M. R. Darragh, S. M. Miller and C. S. Craik, *J. Mol. Biol.*, 2007, **369**, 1041–1051.
21 E. Colombo, A. Desilets, D. Duchene, F. Chagnon, R. Najmanovich, R. Leduc and E. Marsault, *ACS Med. Chem. Lett.*, 2012, **3**, 530–534.
22 O. Avrutina, H. Fittler, B. Glotzbach, H. Kolmar and M. Empting, *Org. Biomol. Chem.*, 2012, **10**, 7753–7762.
23 M. L. J. Korsinczky, H. J. Schirra, K. J. Rosengren, J. West, B. A. Condie, L. Otvos, M. A. Anderson and D. J. Craik, *J. Mol. Biol.*, 2001, **311**, 579–591.
24 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
25 M. J. Costanzo, H. R. Almond, L. R. Hecker, M. R. Schott, S. C. Yabut, H. C. Zhang, P. Andrade-Gordon, T. W. Corcoran, E. C. Giardino, J. A. Kauffman, J. M. Lewis, L. de Garavilla, B. J. Haertlein and B. E. Maryanoff, *J. Med. Chem.*, 2005, **48**, 1984–2008.
26 A. D. Kwong, R. S. Kauffman, P. Hurter and P. Mueller, *Nat. Biotechnol.*, 2011, **29**, 993–1003.
27 K. Hilpert, G. Hansen, H. Wessner, R. Volkmer-Engert and W. Hohne, *J. Biochem.*, 2005, **138**, 383–390.
28 A. Lesner, A. Legowska, M. Wysocka and K. Rolka, *Curr. Pharm. Des.*, 2011, **17**, 4308–4317.
29 A. Legowska, D. Debowski, A. Lesner, M. Wysocka and K. Rolka, *Bioorgan. Med. Chem.*, 2009, **17**, 3302–3307.
30 J. E. Swedberg, S. J. de Veer, K. C. Sit, C. F. Reboul, A. M. Buckle and J. M. Harris, *PLoS One*, 2011, 6.
31 D. Scarpi, J. D. McBride and R. J. Leatherbarrow, *Bioorg. Med. Chem.*, 2004, **12**, 6045–6052.
32 R. G. Boy, W. Mier, E. M. Nothelfer, A. Altmann, M. Eisenhut, H. Kolmar, M. Tomaszowski, S. Kramer and U. Haberkorn, *Mol. Imaging Biol.*, 2010, **12**, 377–385.
33 S. Jiang, P. Li, S. L. Lee, C. Y. Lin, Y. Q. Long, M. D. Johnson, R. B. Dickson and P. P. Roller, *Org. Lett.*, 2007, **9**, 9–12.
34 P. Li, S. Jiang, S. L. Lee, C. Y. Lin, M. D. Johnson, R. B. Dickson, C. J. Michejda and P. P. Roller, *J. Med. Chem.*, 2007, **50**, 5976–5983.
35 Y. Q. Long, S. L. Lee, C. Y. Lin, I. J. Enyedy, S. M. Wang, P. Li, R. B. Dickson and P. P. Roller, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 2515–2519.
36 F. Beliveau, A. Desilets and R. Leduc, *FEBS J.*, 2009, **276**, 2213–2226.
37 J. L. Wike-Hooley, J. Haveman and H. S. Reinhold, *Radiother. Oncol.*, 1984, **2**, 343–366.
38 I. F. Tannock and D. Rotin, *Cancer Res.*, 1989, **49**, 4373–4384.
39 P. Montcourrier, I. Silver, R. Farnoud, I. Bird and H. Rochefort, *Clin. Exp. Metastasis*, 1997, **15**, 382–392.
40 T. Yoshitomi and Y. Nagasaki, *Biointerphases*, 2012, 7, 7.
41 D. Martinez, M. Vermeulen, A. Trevani, A. Ceballos, J. Sabatte, R. Gamberale, M. E. Alvarez, G. Salamone, T. Tanos, O. A. Coso and J. Geffner, *J. Immunol.*, 2006, **176**, 1163–1171.
42 F. Liu, J. E. Park, W. J. Qian, D. Lim, A. Scharow, T. Berg, M. B. Yaffe, K. S. Lee and T. R. Burke Jr., *ChemBioChem*, 2012, **13**, 1291–1296.
43 F. Liu, J. E. Park, W. J. Qian, D. Lim, A. Scharow, T. Berg, M. B. Yaffe, K. S. Lee and T. R. Burke Jr., *ACS Chem. Biol.*, 2012, 7, 805–810.
44 M. Nazare, H. Matter, D. W. Will, M. Wagner, M. Urmann, J. Czech, H. Schreuder, A. Bauer, K. Ritter and V. Wehner, *Angew. Chem., Int. Ed.*, 2012, **51**, 905–911.
45 M. Empting, O. Avrutina, R. Meusinger, S. Fabritz, M. Reinwarth, M. Biesalski, S. Voigt, G. Buntkowsky and H. Kolmar, *Angew. Chem., Int. Ed.*, 2011, **50**, 5207–5211.
46 L. Melendez-Alafort, P. C. Muzzio and A. Rosato, *Anticancer Agents Med. Chem.*, 2012, **12**, 476–499.
47 M. Yoshimoto, T. Hayakawa, M. Mutoh, T. Imai, K. Tsuda, S. Kimura, I. O. Umeda, H. Fujii and K. Wakabayashi, *J. Nucl. Med.*, 2012, **53**, 765–771.
48 A. Kumar, T. Jindal, R. Dutta and R. Kumar, *Ann. Nucl. Med.*, 2009, **23**, 745–751.
49 W. B. Cai, K. Chen, Z. B. Li, S. S. Gambhir and X. Y. Chen, *J. Nucl. Med.*, 2007, **48**, 1862–1870.
50 M. Tischler, D. Nasu, M. Empting, S. Schmelz, D. W. Heinz, P. Rottmann, H. Kolmar, G. Buntkowsky, D. Tietze and O. Avrutina, *Angew. Chem., Int. Ed.*, 2012, **51**, 3708–3712.
51 E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker and K. Karplus, *Proteins*, 2009, 77, 114–122.
52 L. F. Pacios, *Comput. Chem.*, 1997, **21**, 25–34.
53 A. A. Canutescu, A. A. Shelenkov and R. L. Dunbrack, *Protein Sci.*, 2003, **12**, 2001–2014.
54 R. L. Dunbrack and F. E. Cohen, *Protein Sci.*, 1997, **6**, 1661–1681.

2.3. Potent inhibitors of human matriptase-1 based on the scaffold of sunflower trypsin inhibitor

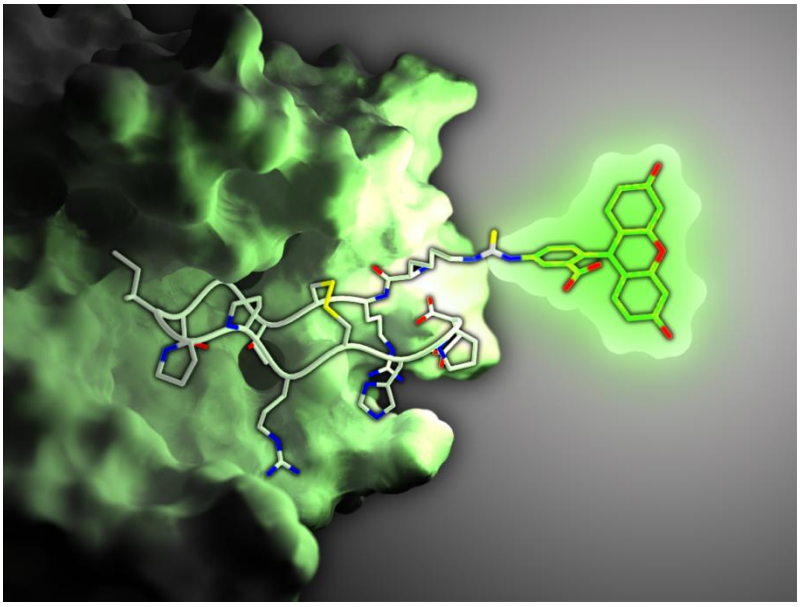
Authors:

Heiko Fittler, Olga Avrutina, Martin Empting, Harald Kolmar

Bibliographic Data:

Journal of Peptide Science,
Volume 20, Issue 6, Page 415-420, 9 May 2014.
DOI: 10.1002/psc.2629
First published online: 10 April, 2014.

Graphical Abstract:



Contributions by H. Fittler

- Developed concept
- Synthesized and isolated all peptidic compounds
- Performed inhibition assays for all compounds
- Wrote the article with O. Avrutina

Research Article

Journal of PeptideScience

Received: 20 January 2014 Revised: 28 February 2014 Accepted: 3 March 2014 Published online in Wiley Online Library: 11 April 2014

(wileyonlinelibrary.com) DOI 10.1002/psc.2629

Potent inhibitors of human matriptase-1 based on the scaffold of sunflower trypsin inhibitor

Heiko Fittler,[a] Olga Avrutina,[a] Martin Empting[b] and Harald Kolmar[a]*

Sunflower trypsin inhibitor-1 (SFTI-1), a bicyclic tetradecapeptide, has become a versatile tool as a scaffold for the development of the inhibitors of therapeutically relevant serine proteases, among them matriptase and kallikreins. Herein, we report the rational design of potent monocyclic and bicyclic inhibitors of human matriptase-1. We found that the presence of positive charge and lack of bulky residues at the peptide *N*-terminus is required for the maintenance of inhibitory activity. Replacement of the *N*-terminal glycine residue by lysine allowed for the chemical conjugation with a fluorophor via the ε-amino group without significant loss of inhibitory activity. Head-to-tail and side-chain-to-tail cyclization resulted in potent inhibitors with comparable activities against matriptase-1. The most potent synthetic bicyclic inhibitor found in this study (K_i = 2.6 nM at pH 7.6) is a truncated version of SFTI-1 (*cyclo*-KRCTKSIPPRCH) lacking a *C*-terminal proline and aspartate residue. It combines an internal disulfide bond with a peptide macrocycle that is formed through side-chain-to-tail cyclization of the ε-amino group of an *N*-terminal lysine and a *C*-terminal proline. Copyright © 2014 European Peptide Society and John Wiley & Sons, Ltd.

Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: rational design; SFTI-1; human matriptase-1; Bowman–Birk inhibitors

Matriptase-1 is a type II transmembrane serine protease that is overexpressed in several tumors and also plays a role in osteoarthritis. It fulfills different functions in diverse epithelial tissues where it is responsible for the development, homeostasis, and cell–cell adhesion [1]. In order to prevent matriptase-1 from being permanently active, it is strictly regulated by the natural inhibitors, hepatocyte growth factor activator inhibitor-1 and 2 [2,3]. Dysregulation of this process leads to the progression of advanced diseases [4]. Indeed, it has been shown that overexpression of matriptase-1 correlates with the development of atherosclerosis [5], osteoarthritis [6], and various kinds of cancer [4,7–10]. Recently, Gao *et al.* demonstrated that matriptase-1 is highly upregulated in chronic lymphocytic leukemia and promotes cancer cell invasion [11]. As a consequence, the development of potent inhibitors of matriptase-1 possessing strong binding characteristics, high selectivity, and reasonable half-life in the human body is a challenging research task.

Several novel matriptase-1 inhibitors have been reported recently, among them small molecules [12,13], cystine-knot miniproteins [14], and peptides from the Bowman–Birk inhibitor (BBI) family. BBIs are peptidic inhibitors of diverse serine proteases isolated from the seeds of legumes and some other plants [15]. The tertiary structure of these molecules having the size of about 6000–9000 Da typically possesses two symmetrical homology domains [16] and is stabilized by seven disulfide bonds. All BBIs share three conserved structural features, namely, a disulfide-bridged macrocycle, an extensive hydrogen network, and a conserved *cis*-Pro at the P3` site [17]. In the case when a BBI inhibitor lacks at least one of these key elements or they are not properly arranged within the structure, its binding activity could be significantly weakened or even lost [18]. As BBI inhibitors are naturally designed to act against diverse proteases, we considered these peptides to serve as a starting point for the molecular design of synthetic matriptase inhibitors.

Among the peptides comprising the pool of BBIs, the sunflower trypsin inhibitor-1 (SFTI-1) is the smallest known member. It has recently attracted closer attention because of its peculiar structural and functional features. Thus, it comprises a backbone of 14 amino acids folded in a cyclic β-sheet that is threaded with an intramolecular disulfide bond [19]. Originally isolated from sunflower seeds, this small and compact oligopeptide is a potent trypsin inhibitor (K_i = 0.1 nM) [20] and has already been used as a scaffold for the design of novel protease inhibitors [21]. Interestingly, the engineered SFTI-based peptides demonstrated potent and selective binding to a variety of serine proteases as matriptase-1, kallikrein-related peptidase, cathepsin B, elastase, chymotrypsin, proteinase K, and β-tryptase [21–28].

Herein, we report the design of novel matriptase-1 inhibitors using the engineered framework of SFTI-1 derived matriptase inhibitor-1 (SDMI-1) as a starting scaffold that was recently reported by our group [29]. SDMI-1 is an open-chain derivative

* *Correspondence to: Harald Kolmar, Clemens-Schöpf Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Alarich-Weiss-Straße 4, 64287 Darmstadt, Germany. E-mail: Kolmar@Biochemie-TUD.de*

a *Clemens-Schöpf Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Alarich-Weiss-Straße 4, 64287 Darmstadt, Germany*

b *Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Department Drug Design and Optimization, Campus C2.3, 66123 Saarbrücken, Germany*

of SFTI-1 possessing a disulfide-bridged 14 amino acid backbone and having two exchanges compared to the wild type, Ile10Arg and Phe12His (Table 1, entry **1** and **2**). It inhibits matriptase-1 with a double-digit nanomolar inhibition constant as determined by the respective *in vitro* assays.

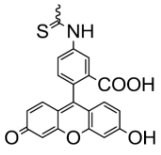
Results and Discussion

Recently, we have reported our study on the development of novel matriptase-1 inhibitors based on the scaffold of SFTI-1 (Table 1, entry **1** and **2**). Stepwise optimization of SFTI-1 resulted in a potent matriptase-1 inhibitor that, compared to the wild-type peptide, displayed two distinguishing features. First, it was found that the open-chain peptide possessing a disulfide moiety but lacking the backbone cyclization motif is more potent compared to the bicyclic compound [30]. Second, amino acid exchanges Ile10Arg and Phe12His led to an improved inhibitory activity of nearly 64-fold compared to the open-chain counterpart SFTI-1[*1,14*]. Indeed, at pH 7.6 K_i of SFTI-1[*1,14*] **1** is 703 nM, while the engineered variant SDMI-1 **2** has a K_i of 11 nM [29].

These finding supposes that the positive charge of the α-amino group contributes favorably to the SDMI-1 binding to matriptase-1. Indeed, the introduction of a bulky fluorescein moiety at

Table 1. Synthesized compounds and their inhibition constants K_i against matriptase, including SFTI-1[*1,14*] (**1**), SDMI-1 (**2**), and SDMI-3 (**4**)

Entry	Structure	K_i/nM
1	*H*-GRCTKSIPPICFPD-*OH*	703 ± 87
2	*H*-GRCTKSIPPRCHPD-*OH*	11.2 ± 2.0
3		328 ± 12
4	*H*-KRCTKSIPPRCHPD-*OH*	2.1 ± 0.1
5		12.2 ± 0.2
6	KRCTKSIPPRCHPD	4.1 ± 0.1
7	GRCTKSIPPRCHPD	3.6 ± 0.1
8	*H*-KRCTKSIPPRCHPD*	4.1 ± 0.2
9	*H*-KRCTKSIPPRCHP*	7.2 ± 0.4
10	*H*-KRCTKSIPPRCH*	2.6 ± 0.2
11	*H*-KRCTKSIPPRCH-*NH*$_2$	8.4 ± 0.2

R^1: 

* Compounds 8-10 are tail-to-side-chain cyclized

Journal of PeptideScience

the α-amine of Gly1 through the β-alanine linker caused dramatic decrease of the inhibitory potency: the resulting peptide 3 (Table 1) demonstrated a very moderate K_i of 328 nM against matriptase-1. This outcome is supported by the recent results of Lesner and co-workers who showed that monocyclic variants of SFTI-1[*1,14*] lacking the Arg2 and labeled *N*-terminally with fluorescent derivatives of benzoxazole showed lower association constants against trypsin (one or two orders magnitude, depending on the used derivative) compared to the parent wild-type peptide [31]. Leaping ahead, it is important to mention that in the case when the same moiety was introduced via the side chain of Lys1, thus preserving the charge at the *N*-terminus, the binding to the target of the resulting peptide 5 (Table 1) was significantly enhanced. To gain further insight, we conducted a molecular modeling study using the published X-ray structure of matriptase-1 in complex with SFTI-1 (Protein Data Bank entry 3P8F) [32]. The modeled structures of compounds 3 and 5 provided further evidence for a favorable interaction between the *N*-terminus of 5 and surface exposed residues of matriptase-1 (Figure 1(A)). In detail, the proximal carboxy function of the Asp217 side chain may form a salt bridge with the free amine, which is not available in the *N*-terminally functionalized derivative 3 (Figure 1(A)).

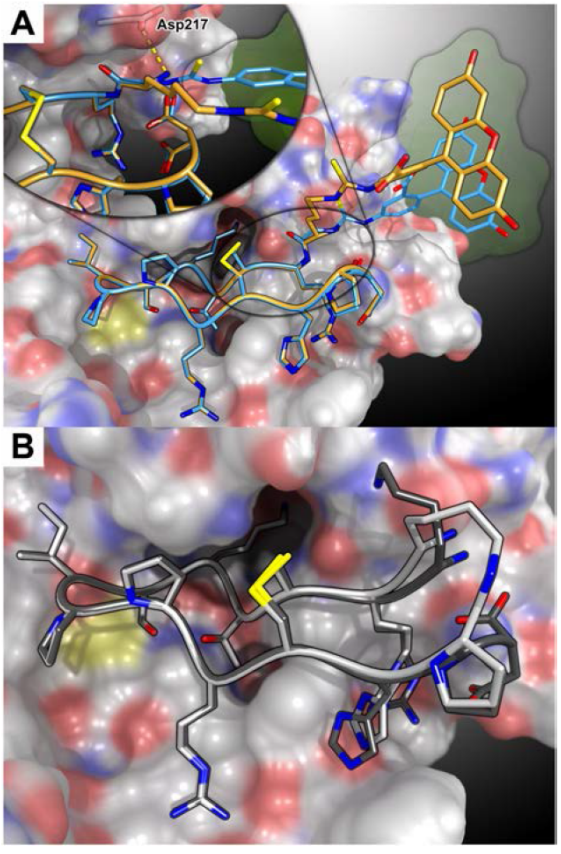


Figure 1. *In silico* models of peptides (A) 3 (light blue carbons) and 5 (orange carbons) as well as (B) 4 (grey carbons) and 9 (white carbons) in complex with matriptase-1 (surface). Peptide backbones are depicted as tubes, side chains as sticks, and fluorophor moieties are highlighted (green, transparent surface). A possible salt bridge between Asp217 and the *N*-terminus of 5 is shown as yellow dashed sticks. Protein carbons, white; nitrogen, blue; oxygen, red; sulfur, yellow; and hydrogen left out for clarity.

Interestingly, an SDMI variant 4 that has a lysine residue at position 1 displayed a five-fold enhanced inhibitory activity compared to SDMI variant 2 corroborating the notion that a positively charged environment is required at the SDMI *N*-terminus. However, variant 7 that is head-to-tail cyclized through the *N*-terminal glycine and, therefore, lacks the charge at that position surprisingly demonstrated a single-digit nanomolar inhibition constant. This finding suggests that rigid, compact architecture of the *N*-terminal region that excludes conformational clash and allows for the perfect match with the matriptase-1 binding pocket is probably more important than the charge.

Introduction of a lysine residue offered the possibility for ring closure or *N*-terminal modification via the lysine ε-amino group while maintaining the α-amino group. To that end, the commercially available Boc-Lys(Fmoc)-OH building block was used. Its base-labile fluorenylmethoxycarbonyl protecting group can be cleaved on-support thus providing the free ε-amine function for further conjugation, while the *N*-terminal Boc protecting group is removed after the chain assembly has been completed. Therefore, every desired modification could be performed at the side chain leaving the *N*-terminus intact.

As mentioned earlier, compound 5 bearing the bulky fluorescein coupled to the side chain of residue Lys1 (Figure 1(A)) via a β-alanine linker was 27 times more potent compared to compound 3, where this moiety was introduced through the *N*-terminal amine (K_i of 12.2 and 328 nM, respectively) confirming the requirement of positive charge in this region of the inhibitor. On the other hand, the bulky substituents at the *N*-terminus might additionally clash with the surface of matriptase-1 because of steric hindrance, thus leading to further decreased binding potency.

In order to study the influence of backbone cyclization on bioactivity, peptide 4 was head-to-tail or tail-to-side chain cyclized using either the α-amine or ε-amine of Lys1 leading to compounds 6 and 8, respectively. The resulting bicyclic molecules showed similar inhibition constants against matriptase-1 compared to the monocyclic parent SDMI-3 (4).

Summing up, positive charge at the SDMI *N*-terminal region is important for the binding with matriptase-1, but substituents at this position also play a significant role as they have to be compact or flexible enough to match with the size of the matriptase's binding pocket. On a long term, it would be desirable to install functional groups in SDMI that allow for its bio-orthogonal conjugation with, e.g. an antibody Fc fragment, aimed at achieving a prolonged serum half-life of the inhibitor. To this end, the optimal position at the SDMI-1 site has to be identified where the conjugation would not affect the biological activity of the resulting construct. Indeed, it has been shown that most of amino acid residues of the SFTI backbone are either crucial for the interaction with the surface of the protease, or not suitable for conjugation due to the structural requirements of the substrate binding site leading to steric hindrance [33,34]. As a consequence, we decided to modify the SDMI-1 *N*-terminal region taking into consideration the requirement to maintain its positive charge and to avoid bulky substituents [29]. Hence, the lysine side chain seems the position of choice for further bioconjugations.

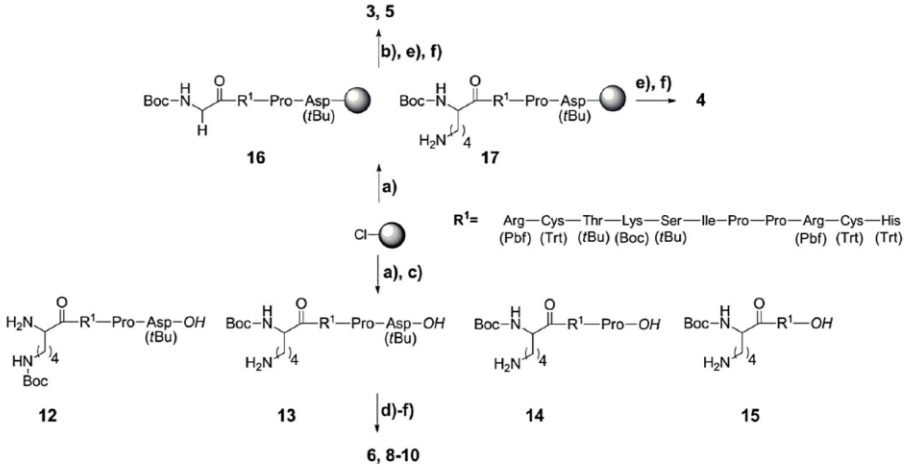


Figure 2. (a) Synthesis of the precursor molecules **12–17** via Fmoc-SPPS. (b) Coupling of the fluorophore with 2 eq. fluorescein isothiocyanate (FITC) and 4 eq. DIEA in DMF. (c) Mild cleavage from the resin using acetic acid/DCM/MeOH (5 : 4 : 1; v : v : v) left all acid-labile protecting groups intact. (d) Macrocyclization of the free amine and the carboxylic group of the *C*-terminus with HOBt (5 eq.), PyBop (5 eq.), and DIEA (10 eq.) in DMF overnight at ambient temperature. (e) Acidic deprotection using TFA/H_2O/anisole/TES (47 : 1 : 1 : 1; v : v : v : v). (f) Air-mediated oxidative macrocyclization in 100 mM $(NH_4)_2CO_3$ aq. (pH = 8.4, 1 mg ml^{-1} peptide). Compounds **3–10** were isolated via preparative reversed-phase HPLC.

Additionally, we synthesized the truncated bicyclic variants **9** and **10** lacking either the *C*-terminal aspartic acid or the Pro-Asp sequence, respectively. The inhibitory potency of compounds **7–11** was of the same range (Table 1), indicating that the last two *C*-terminal residues only make a minor contribution to matriptase-1 binding. Indeed, upon inhibition mainly the substrate-mimicking loop of SFTI-based peptides directly interacts with the surface of an enzyme (Figure 1(B)). Compound **10** that lacks the *C*-terminal proline and aspartate residue and contains a tail-to-side-chain cyclization is a potent and tight-binding inhibitor of matriptase-1. Therefore, in view of further applications (considering, e.g. plasma half-life and proteolytic stability), it is an important finding that the second macrocycle could be achieved without the dramatic loss of inhibitory activity.

Conclusion

Oligopeptide SFTI-1 is a versatile scaffold enabling the design of novel protease inhibitors by simple grafting of the respective substrate sequence in its inhibitor loop. In this study, we developed a number of highly potent inhibitors of human matriptase-1 based on the engineered SFTI variant, the open-chain disulfide SDMI-1. We found that the substitution of the *N*-terminal glycine by lysine led to two important advantages. First, it resulted in a potent matriptase-1 binder (SDMI-3, **4**) with an inhibition constant in the low nanomolar range. Second, it provided an additional addressable site ensuring the introduction of tailor-made functionalities. Furthermore, we studied the influence of the cyclic backbone on bioactivity and found that the bicyclic variants, both full-length and truncated, provide inhibitory strength in the same range, indicating that the *C*-terminal amino acids are not required for the efficient interaction with the surface of the target enzyme. The activity of this bicyclic minimized peptide against matriptase-1 was conserved with interesting ramifications for further applications *in vivo*.

General

All chemicals and solvents were used as supplied by Iris Biotech (Marktredwitz, Germany), Bachem (Bubendorf, Switzerland), Novabiochem (Darmstadt, Germany), Sigma-Aldrich (Steinheim, Germany), Roth (Karlsruhe, Germany), and Varian (Agilent, Böblingen, Germany).

Peptides were analyzed on an analytical RP-HPLC from Varian, model 920-LC using a Phenomenex Hypersil 5u BDS C18 LC column (150 × 4.6 mm, 5 µm, 13 nm), and isolated by semi-preparative RP-HPLC (Varian) using a Phenomenex Luna 5u C18 LC column (250 × 12.2 mm, 5 µm, 10 nm) as stationary phase. Both eluents A (water) and B (90% aq. acetonitrile) contained 0.1% aq. TFA.

Electrospray ionization mass spectra were collected on a Shimadzu LCMS-2020 equipped with a Phenomenex Jupiter 5u C4 LC column (50 × 1 mm, 5 µm, 30 nm) using an eluent system of 0.1% aq. formic acid, LC-MS grade (eluent A) and acetonitrile containing 0.1% aq. formic acid, LC-MS grade (eluent B) (see Figures S1–S9 and Table S1 in Supporting Information).

Chemistry

Peptides **3–10** were synthesized on a 2-chlorotrityl chloride resin (Iris Biotech) by standard Fmoc-SPPS with subsequent cyclization towards either monocyclic or bicyclic compounds (Figure 2).

The first amino acid was coupled by adding a solution of 103 mg Fmoc-L-Asp(tBu)-OH (0.25 mmol) and 34 µl DIEA (1 mmol) in a minimal amount of DCM to the pre-swollen resin. The mixture was gently shaken for 2 h at ambient temperature. After filtration of the coupling solution, the resin was washed with DCM-methanol-DIEA (17 : 2 : 1; 3 × 5 ml), dimethylformamide (3 × 5 ml), and DCM (3 × 5 ml).

Peptide **11** was synthesized on an AmphiSpheres 40 RAM resin (0.44 mmol g^{-1}) on a 0.05 mmol scale.

Journal of PeptideScience

The remaining amino acids were coupled manually by double coupling using 4 eq. of the corresponding amino acid, 3.9 eq. of 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and 8 eq. of DIEA. For the couplings of histidine and cysteine, 6 eq. of 2,4,4-trimethylpyridine (collidine) were used instead of DIEA. The attachment was achieved under microwave assistance at 50 °C, 25 W for 20 min. Fmoc deprotection of the *N*-terminal amino acid was conducted using 20% piperidine in DMF at 50 °C, 25 W for 5 min (2 × 5 ml).

The introduction of the fluorophore required the coupling of *β*-alanine as a linker to prevent removal of the *N*-terminal amino acid [35], which was installed following the procedure described earlier. The solution of 2 eq. FITC and 4 eq. DIEA in DMF was added to the resin, and the coupling was conducted overnight at ambient temperature.

All peptides were cleaved by a standard cleavage cocktail of TFA/anisole/triethylsilane/H_2O (47 : 1 : 1 : 1; v : v : v : v) and dithiotreitol to suppress unwanted oxidation of cysteines. After 2 h, the peptides were precipitated in cold diethylether and washed afterwards again with diethylether (2 × 40 ml).

The cyclic peptides were obtained by acidolytic cleavage under mild conditions with acetic acid/DCM/methanol (5 : 4 : 1; v : v : v) for 2 h. The solvent was evaporated, and the resulting yellow oil was washed with *n*-hexane (5 × 5 ml). The residue was dissolved in H_2O/acetonitrile (1 : 1), and the peptide was isolated by freeze-drying. Then, the protected peptides were dissolved in DMF and 5 eq. HOBt, 5 eq. benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBop) as well as 10 eq. DIEA were added to the resulted mixture. The cyclization was conducted at ambient temperature overnight. Finally, the solvent was evaporated under reduced pressure, and protecting groups were removed applying the standard cleavage cocktail.

The crude peptides were oxidized in 100 mM $(NH_4)_2CO_3$ aq. (1 mg ml^{-1}, pH = 8.4) monitoring the oxidative folding by RP-HPLC. Afterwards, the solvent was removed by freeze-drying, and the peptides were isolated by semi-preparative RP-HPLC (see Figures S10–S12 and Table S2 in Supporting Information). This resulted in the following yields for compounds **3–11**. **3**, 9.2 mg (8.8%); **4**, 15.3 mg (18.7%); **5**, 10.3 mg (10.2%); **6**, 6.4 mg (8.2%); **7**, 7.8 mg (10.1%); **8**, 13.3 mg (16.4%); **9**, 0.7 mg (1.0%); **10**, 3.6 mg (4.8%); **11**, 8.3 mg (11.7%).

Enzyme Production

Matriptase-1 was recombinantly produced as reported previously [30]. *In brevi*, the catalytic domain of human matriptase-1 was overexpressed in *Escherichia coli* BL21-CodonPlus(DE3)-RP cells (Stratagene) bearing the expression vector pET42dest-His-hMatl. The resulting inclusion bodies comprising the unfolded enzyme were purified and folded. Subsequent autoproteolytic activation resulted in the functional catalytic domain.

In Vitro Assays

All experiments were performed in triplicate. The respective inhibitor was pre-equilibrated in various concentrations at pH 7.6 for 30 min with matriptase-1 (final concentration: [E] = 0.9 nM). Afterwards, the chromogenic substrate Boc-QAR-*p*NA was added, and the absorption was measured at 405 nm in 60 s intervals over 30 min at ambient temperature using the Tecan GENios microplate reader [29]. The resulting dose–response curves were fitted by Eqn (1) (see Figures S13–S16 in Supporting Information).

$$\frac{v}{v_0} = \left(1 + \left(\frac{[I]}{IC_{50}}\right)^{Hill}\right)^{-1} \quad (1)$$

Substrate-independent inhibition constants K_i were calculated from respective IC_{50} values using Eqn (2).

$$K_i = IC_{50}\left(1 + \frac{[S]}{K_m}\right)^{-1} \quad (2)$$

Molecular Modeling

In silico structures shown in Figure 1 were modeled on the basis of published X-ray structure of matriptase-1 in complex with SFTI-1 (PDB-ID: 3P8F) using the YASARA *structure package* (YASARA Biosciences GmbH, Vienna, Austria) and the AMBER03 force field [36]. First, amino acids of the ligand were exchanged to resemble the sequence of the respective compounds. Then, optimal side chain rotamers were calculated using the built-in optimize command [37]. Non-natural moieties within inhibitors **3**, **5**, and **9** (the fluorophores and tail-to-side chain macrocyclization motif) were modeled manually into the ligand:receptor complex. Then, resulting structures where energy minimized with fixed (constrained) receptor atoms. (Persistence of Vision Raytracer).

Acknowledgement

This work was supported in part by the Deutsche Forschungsgemeinschaft SPP1623 through grant KO1390/10-1.

References

1 Miller GS, List K. The matriptase-prostasin proteolytic cascade in epithelial development and pathology. *Cell Tissue Res.* 2013; **351**: 245–253.
2 Oberst MD, Chen LY, Kiyomiya K, Williams CA, Lee MS, Johnson MD, Dickson RB, Lin CY. HAI-1 regulates activation and expression of matriptase, a membrane-bound serine protease. *Am. J. Physiol. Cell Physiol.* 2005; **289**: C462–470.
3 Szabo R, Hobson JP, List K, Molinolo A, Lin CY, Bugge TH. Potent inhibition and global co-localization implicate the transmembrane Kunitz-type serine protease inhibitor hepatocyte growth factor activator inhibitor-2 in the regulation of epithelial matriptase activity. *J. Biol. Chem.* 2008; **283**: 29495–29504.
4 Oberst MD, Johnson MD, Dickson RB, Lin CY, Singh B, Stewart M, Williams A, al-Nafussi A, Smyth JF, Gabra H, Sellar GC. Expression of the serine protease matriptase and its inhibitor HAI-1 in epithelial ovarian cancer: correlation with clinical outcome and tumor clinicopathological parameters. *Clin. Cancer Res.* 2002; **8**: 1101–1107.
5 Seitz I, Hess S, Schulz H, Eckl R, Busch G, Montens HP, Brandl R, Seidl S, Schomig A, Ott I. Membrane-type serine protease-1/matriptase induces interleukin-6 and -8 in endothelial cells by activation of protease-activated receptor-2: potential implications in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* 2007; **27**: 769–775.
6 Milner JM, Patel A, Davidson RK, Swingler TE, Desilets A, Young DA, Kelso EB, Donell ST, Cawston TE, Clark IM, Ferrell WR, Plevin R, Lockhart JC, Leduc R, Rowan AD. Matriptase is a novel initiator of cartilage matrix degradation in osteoarthritis. *Arthritis Rheum.* 2010; **62**: 1955–1966.
7 Kang JY, Dolled-Filhart M, Ocal IT, Singh B, Lin CY, Dickson RB, Rimm DL, Camp RL. Tissue microarray analysis of hepatocyte growth factor/Met pathway components reveals a role for Met, matriptase, and hepatocyte growth factor activator inhibitor 1 in the progression of node-negative breast cancer. *Cancer Res.* 2003; **63**: 1101–1105.
8 Saleem M, Adhami VM, Zhong W, Longley BJ, Lin CY, Dickson RB, Reagan-Shaw S, Jarrard DF, Mukhtar H. A novel biomarker for staging human prostate adenocarcinoma: overexpression of matriptase with concomitant loss of its inhibitor, hepatocyte growth factor activator inhibitor-1. *Cancer Epidemiol. Biomarkers Prev.* 2006; **15**: 217–227.

9 Santin AD, Cane S, Bellone S, Bignotti E, Palmieri M, De Las Casas LE, Anfossi S, Roman JJ, O'Brien T, Pecorelli S. The novel serine protease tumor-associated differentially expressed gene-15 (matriptase/MT-SP1) is highly overexpressed in cervical carcinoma. *Cancer* 2003; 98: 1898–1904.
10 Lee JW, Yong Song S, Choi JJ, Lee SJ, Kim BG, Park CS, Lee JH, Lin CY, Dickson RB, Bae DS. Increased expression of matriptase is associated with histopathologic grades of cervical neoplasia. *Hum. Pathol.* 2005; 36: 626–633.
11 Gao L, Liu M, Dong N, Jiang Y, Lin CY, Huang M, Wu D, Wu Q. Matriptase is highly upregulated in chronic lymphocytic leukemia and promotes cancer cell invasion. *Leukemia* 2013; 27: 1191–1194.
12 Colombo E, Desilets A, Duchene D, Chagnon F, Najmanovich R, Leduc R, Marsault E. Design and synthesis of potent, selective inhibitors of matriptase. *Acs Medicinal Chemistry Letters* 2012; 3: 530–534.
13 Uhland K, Siphos B, Arkona C, Schuster M, Petri B, Steinmetzer P, Mueller F, Schweinitz A, Steinmetzer T, Van De Locht A. Use of IHC and newly designed matriptase inhibitors to elucidate the role of matriptase in pancreatic ductal adenocarcinoma. *Int. J. Oncol.* 2009; 35: 347–357.
14 Glotzbach B, Reinwarth M, Weber N, Fabritz S, Tomaszowski M, Fittler H, Christmann A, Avrutina O, Kolmar H. Combinatorial optimization of cystine-knot peptides towards high-affinity inhibitors of human matriptase-1. *Plos One* 2013; 8: e76956.
15 Laskowski Jr M, Kato I. Protein inhibitors of proteinases. *Annu. Rev. Biochem.* 1980; 49: 593–626.
16 Jaulent AM, Brauer AB, Matthews SJ, Leatherbarrow RJ. Solution structure of a novel C2-symmetrical bifunctional bicyclic inhibitor based on SFTI-1. *J. Biomol. NMR* 2005; 33: 57–62.
17 Daly NL, Chen YK, Foley FM, Bansal PS, Bharathi R, Clark RJ, Sommerhoff CP, Craik DJ. The absolute structural requirement for a proline in the P3'-position of Bowman-Birk protease inhibitors is surmounted in the minimized SFTI-1 scaffold. *J. Biol. Chem.* 2006; 281: 23668–23675.
18 Tischler M, Nasu D, Empting M, Schmelz S, Heinz DW, Rottmann P, Kolmar H, Buntkowsky G, Tietze D, Avrutina O. Braces for the peptide backbone: insights into structure-activity relationships of protease inhibitor mimics with locked amide conformations. *Angew. Chem. Int. Ed. Engl.* 2012; 51: 3708–3712.
19 Debowski D, Lukajtis R, Filipowicz M, Strzelecka P, Wysocka M, Legowska A, Lesner A, Rolka K. Hybrid analogues of SFTI-1 modified in P1 position by beta- and gamma-amino acids and N-substituted beta-alanines. *Biopolymers* 2013; 100: 154–159.
20 Luckett S, Garcia RS, Barker JJ, Konarev AV, Shewry PR, Clarke AR, Brady RL. High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Biol.* 1999; 290: 525–533.
21 Fittler H, Avrutina O, Glotzbach B, Empting M, Kolmar H. Combinatorial tuning of peptidic drug candidates: high-affinity matriptase inhibitors through incremental structure-guided optimization. *Org. Biomol. Chem.* 2013; 11: 1848–1857.
22 Quimbar P, Malik U, Sommerhoff CP, Kaas Q, Chan LY, Huang YH, Grundhuber M, Dunse K, Craik DJ, Anderson MA, Daly NL. High-affinity cyclic peptide matriptase inhibitors. *J. Biol. Chem.* 2013; 288: 13885–13896.
23 Rafal L, Anna L, Magdalena W, Dawid D, Adam L, Krzysztof R. Analogues of trypsin inhibitor SFTI-1 modified in the conserved P(1)' position by synthetic or non-proteinogenic amino acids retain their inhibitory activity. *J. Pept. Sci.* 2011; 17: 281–287.
24 McBride JD, Freeman HN, Leatherbarrow RJ. Selection of human elastase inhibitors from a conformationally constrained combinatorial peptide library. *Eur. J. Biochem.* 1999; 266: 403–412.
25 Swedberg JE, de Veer SJ, Sit KC, Reboul CF, Buckle AM, Harris JM. Mastering the canonical loop of serine protease inhibitors: enhancing potency by optimising the internal hydrogen bond network. *Plos One* 2011; 6: e19302.
26 Hilpert K, Hansen G, Wessner H, Volkmer-Engert R, Hohne W. Complete substitutional analysis of a sunflower trypsin inhibitor with different serine proteases. *J. Biochem.* 2005; 138: 383–390.
27 Legowska A, Debowski D, Lesner A, Wysocka M, Rolka K. Introduction of non-natural amino acid residues into the substrate-specific P1 position of trypsin inhibitor SFTI-1 yields potent chymotrypsin and cathepsin G inhibitors. *Bioorg. Med. Chem.* 2009; 17: 3302–3307.
28 Rothemund S, Sonnichsen FD, Polte T. Therapeutic potential of the peptide leucine arginine as a new nonplant Bowman-Birk-like serine protease inhibitor. *J. Med. Chem.* 2013; 56: 6732–6744.
29 Fittler H, Avrutina O, Glotzbach B, Empting M, Kolmar H. Combinatorial tuning of peptidic drug candidates: high-affinity matriptase inhibitors through incremental structure-guided optimization. *Organic & biomolecular chemistry* 2013; 11: 1848–1857.
30 Avrutina O, Fittler H, Glotzbach B, Kolmar H, Empting M. Between two worlds: a comparative study on in vitro and in silico inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH. *Organic biomolecular chemistry* 2012; 10: 7753–7762.
31 Lesner A, Karna N, Psurski M, Legowska A, Wysocka M, Guzow K, Sieradzan A, Sienczyk M, Trzonkowski P, Pikula M, Zielinski M, Kosikowska P, Lukajtis R, Legowska M, Debowski D, Wiczk W, Rolka K. Fluorescent analogs of trypsin inhibitor SFTI-1 isolated from sunflower seeds - synthesis and applications. *Biopolymers* 2014; 102: 124–135.
32 Yuan C, Chen L, Meehan EJ, Daly N, Craik DJ, Huang M, Ngo JC. Structure of catalytic domain of Matriptase in complex with Sunflower trypsin inhibitor-1. *Bmc Struct. Biol.* 2011; 11: 30.
33 Quimbar P, Malik U, Sommerhoff CP, Kaas Q, Chan LY, Huang Y-H, Grundhuber M, Dunse K, Craik DJ, Anderson Ma, Daly NL. High-affinity cyclic peptide matriptase inhibitors. *The Journal of biological chemistry* 2013; 288: 13885–13896.
34 Li P, Jiang S, Lee S-L, Lin CY, Johnson MD, Dickson RB, Michejda CJ, Roller PP. Design and synthesis of novel and potent inhibitors of the type II transmembrane serine protease, matriptase, based upon the sunflower trypsin inhibitor-1. *J. Med. Chem.* 2007; 50: 5976–5983.
35 Jullian M, Hernandez A, Maurras A, Puget K, Amblard M, Martinez J, Subra G. N-terminus FITC labeling of peptides on solid support: the truth behind the spacer. *Tetrahedron Lett.* 2009; 50: 260–263.
36 Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* 2003; 24: 1999–2012.
37 Krieger E, Joo K, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins* 2009; 77(Suppl 9): 114–122.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site.

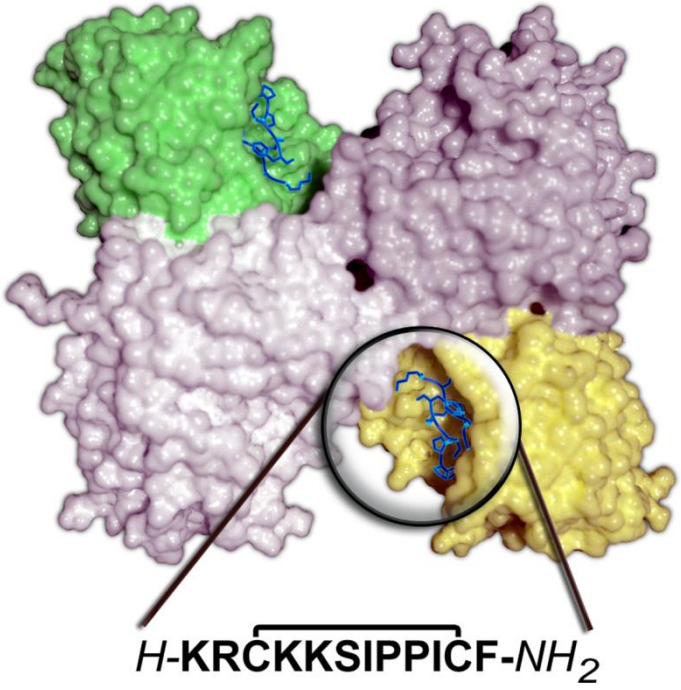
2.4. Engineering a Constrained Peptidic Scaffold towards Potent and Selective Furin Inhibitors.

Authors:
Heiko Fittler, Alexander Depp, Olga Avrutina, Sven O. Dahms, Manuel Than, Martin Empting, Harald Kolmar

Bibliographic Data:

ChemBioChem
Volume 16, Issue 17, Page 2441-2444, 23 November 2015.
DOI: 10.1002/cbic.201500447
First published online: 19 October, 2015.

Graphical Abstract:



Contributions by H. Fittler

- Developed concept
- Synthesized and isolated all peptidic compounds
- Performed inhibition assays for all compounds
- Wrote the article with O. Avrutina

ChemPubSoc Europe

DOI: 10.1002/cbic.201500447

CHEMBIOCHEM Communications

Engineering a Constrained Peptidic Scaffold towards Potent and Selective Furin Inhibitors

Heiko Fittler,[a] Alexander Depp,[a] Olga Avrutina,[a] Sven O. Dahms,[b] Manuel E. Than,[b] Martin Empting,[c] and Harald Kolmar*[a]

We report the engineering of the monocyclic sunflower trypsin inhibitor (SFTI-1[1,14]) into a potent furin inhibitor. In a rational approach, we converted the native scaffold of this trypsin-like serine protease inhibitor into a subtilisin-like one by substitutions in the canonical and, particularly, in the substrate-binding loop. Although the substrate sequence for furin is Arg-X-Arg/Lys-Arg↓, the most potent inhibitor had a lysine at position P1. C-terminally truncated versions demonstrated the strongest activity, thus suggesting a lack of interaction between this motif and the surface of furin. This observation was further supported by molecular modeling. With an inhibition constant of 0.49 nM, the engineered peptide H-KRCKKSIPPICF-NH_2 is a promising compound for further development of furin inhibitors aimed at controlling the activity of this protease in vitro and in vivo.

Furin, a subtilisin-like type-I transmembrane serine protease, is ubiquitously expressed in all vertebrates and many invertebrates. It belongs to the family of the proprotein convertases (PCs).[1–2] Discovered in 1986 and later classified as an endoprotease, it was the first identified PC.[3–4] This calcium-dependent serine protease is structurally related to bacterial subtilisin and yeast kexin,[5] and is expressed as a proprotein with a cleavable prodomain (83 amino acids) that functions as an intramolecular chaperone for activation and folding of the active enzyme.[6] It is mainly located in the trans-Golgi network and is responsible for the maturation of a vast number of protein precursors: cell-surface proteins, zymogens, hormones, and receptor proforms.[7] It can also inactivate certain proteins, such as growth factors, receptors, adhesion molecules, metalloproteinases, viral glycoproteins, and bacterial toxins.[8] Roebroek et al. demonstrated the importance of furin in a homozygous knock-out mouse model: embryos lacking this enzyme died before the second week.[9] As furin is involved in various diseases, such as atherosclerosis, Alzheimer's disease, inflammatory disorders, cancer, and, particularly, long-term pathogen infections, development of potent inhibitors has attracted keen research attention.[10–12]

One of the first reported inhibitors of this PC was the protein-derived compound α_1-antitrypsin Portland (α_1-PDX; K_i = 600 nM).[13] It was engineered by substituting the native AIPM amino acid motif to a consensus furin cleavage site RIPR. Smaller molecules were subsequently found to inhibit furin, among them oligoarginine peptides,[14] with nona-L-arginine (K_i = 40 nM)[15] and nona-D-arginine (K_i = 1.3 nM) being the most potent.[16] Later, irreversible peptidyl chloromethylketones were discovered,[17–18] and Jiao et al. developed strong and reversible inhibitors based on 2,5-dideoxystreptamine (K_i = 6 nM).[19]

Libraries of small molecules comprising a 4-aminomethylbenzamidine group at position P1 were reported by Becker et al;[20] the most potent inhibitor was decorated with a phenylacetyl group at the N terminus (K_i = 810 pM). Optimization of this compound was conducted by screening for a new chemical moiety at position P5. The introduction of a positively charged residue at the N terminus and a (4-guanidino-methyl)-phenylacetyl group at position P5 resulted in a variant with a K_i of 8 pM.[21] The mode of binding of *meta*-guanidinomethyl-Phac-RVR-Amba to furin was elucidated by co-crystallization (PDB ID: 4OMC).[22]

Despite the importance of this protease in a vast number of diseases, research of furin inhibitors is still limited, and novel candidates combining strong potency with good selectivity are needed. Here, we report the development of new peptidic furin inhibitors based on the constrained scaffold of the monocyclic version of sunflower trypsin inhibitor, SFTI-1[1,14]. In a previous investigation, we demonstrated the versatility of this smallest Bowman-Birk inhibitor (BBI) in structure–activity studies, and for development of potent inhibitors of serine proteases such as human matriptase-1, by stepwise substitutions of singular residues. The most promising candidate, SDMI-3 (an SFTI-derived matriptase inhibitor), displayed a K_i of 2.1 nM, which is over 300-fold more potent than for the wild-type inhibitor SFTI-1[1,14].[23]

In this study, we chose the SFTI scaffold as a starting point to develop new furin inhibitors, as it is constrained, synthetically accessible, and its bioactivity directly reflects the changes within the backbone.[24] Two scaffolds were used: the monocyclic peptide SFTI-1[1,14] **1** and the engineered variant SDMI-3 **2** (Table 1).[23] Against furin, **1** showed K_i = 3.5 μM, whereas **2** already revealed K_i = 24 nM, so was a good starting point for fur-

[a] H. Fittler, A. Depp, Dr. O. Avrutina, Prof. Dr. H. Kolmar
Clemens-Schöpf-Institut für organische und Biochemie
Technische Universität Darmstadt
Alarich-Weiss Strasse 4, 64287 Darmstadt (Germany)
E-mail: kolmar@biochemie-tud.de

[b] Dr. S. O. Dahms, Dr. M. E. Than
Leibniz Institute for Age Research–Fritz Lipmann Institute (FLI)
Beutenbergstrasse 11, 07745 Jena (Germany)

[c] Dr. M. Empting
Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS)
Campus C2.3, 66123 Saarbrücken (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201500447.

ther stepwise optimization. Compared to **1**, **2** has three backbone substitutions, Gly1Lys at the canonical loop, and Ile10Arg and Phe12His in the C-terminal β-strand. Therefore, the role of these structural elements for furin inhibition was further investigated in this study.

As the crystal structure of furin (PDB ID: 1P8J)[25] indicates a highly negative charge around its active site (see also Figure 1),[22] it is not surprising that (positively charged) oligoarginines demonstrated strong binding to its surface, and hence good inhibition.[14] Thus, the first step was the implementation of the natural furin substrate sequence (Arg-X-Arg/Lys-Arg↓; arrow indicates the cleavage site) into the canonical loop of **1**. We expected an improvement in activity, compared to **2**, because Arg at position P1 was favored over Lys.[26,27] Surprisingly, the resulting peptide, **3**, had dramatically impaired activity (K_i ≈1.2 μM). Our modeling suggested that the privileged conformation of the BBI loop enables the Lys residue at P1 to penetrate deep enough into the S1 site to facilitate favorable contacts with Asp258 and Asp306; an arginine side chain attached might reach too far into this pocket thus causing steric clashes at the bottom.

Hence, we retained Lys at P1 and performed other mutations within the backbone of **3**: **4** (Lys at P1, Arg at P2; K_i= 29.3 nM), was significantly improved over **3**; inhibition by **5** (Gly1Lys; K_i=8.8 nM) and **6** (Gly1Arg; K_i=9.1 nM) suggested the importance of a positively charged side chain for the N-terminal residue.

To study the contribution of the residues 10 and 12 to furin binding, we exchanged Arg10 (as in **2**) to Ile, and His12 to Phe (as in **1**). These manipulations (peptide **7**) did not alter the performance compared to **5**, despite the change of amino acids at the C-terminal β-strand (K_i=9.8 nM). This indicates that the interaction between this structural element and furin's surface is not very pronounced, as supported by our in silico model, as few interactions between residues 12–14 were detected (Figure S1 in the Supporting Information).

In order to test the requirement for the guanidine moiety of Arg at P4 for the favorable interaction with the surface of furin, we substituted it to Ala. Indeed, the resulting construct (**8**) showed only weak inhibitory activity (K_i=1.3 μM).

Parent scaffold SFTI-1[1,14] contains a *cis*-proline; this in additions to an extensive hydrogen network and a disulfide-bridged macrocycle is a signature of Bowman–Birk inhibitors.[29] In order to examine the role of this structural element, we exchanged the *cis*-Pro8 to Ala (**9**) to obtain a full *trans* amide backbone. As expected, this construct was a very poor furin inhibitor (K_i≈2 μM).

Furin tolerates either Lys or Arg at P2 of its substrate sequence. Comparison of **5** (Arg4) and **10** (Lys4) showed that the lysine-comprising variant was slightly more potent (K_i=8.8 and 3.8 nM, respectively). Therefore, Lys was used at P2 in peptides **11–14**, **16**, and **17**.

In our previous study we demonstrated a beneficial contribution of C-terminal backbone truncation on the inhibition of matriptase-1.[30] Likewise, we synthesized peptides shortened by two (**11**, **14**) or three (**12**, **13**) C-terminal residues. Additionally, in order to eliminate a negative charge at the carboxy terminus (in order to prevent repulsion with furin's surface) the latter was converted into a neutral amide upon cleavage from Rink amide resin. Interestingly, these modifications had a pronounced effect on the activity: sub-nanomolar K_i values for **11** and **12** (490 and 710 pM, respectively). Taken together, our data suggest that the carboxy terminal section of the SFTI derivative does not contribute to binding with furin and, as a consequence, the peptide can be truncated to Cys11, which is essential for the disulfide-bridged scaffold.

The substitutions in constructs **5** and **7** at position 10 were similarly accomplished in the truncated variants **13** (Arg10) and **14** (Ala10). Both revealed good inhibition of furin (single-digit nanomolar K_i values: 4.7 and 3.7 nM, respectively), although the isoleucine side chain at position 10 (**11** and **12**) resulted in even higher activities.

Table 1. Synthesized inhibitors and their inhibition constants against furin. Residue numbering according to Schechter and Berger.[28]

	P5	P4	P3	P2	P1	P1′–P4′						K_i [nM]	Error [nM]
	1	2	3	4	5	6–9	10	11	12	13–14	C terminus		
1	G	R	C	T	K	SIPP	I	C	F	PD	OH	35234	1579
2	**K**	R	C	T	K	SIPP	R	C	H	PD	OH	24.1	0.8
3	**G**	R	C	**R**	**R**	SIPP	R	C	H	PD	OH	1157	10
4	**G**	R	C	**R**	K	SIPP	R	C	H	PD	OH	29.3	0.6
5	**K**	R	C	**R**	K	SIPP	R	C	H	PD	OH	8.8	0.5
6	**R**	R	C	**R**	K	SIPP	R	C	H	PD	OH	9.1	0.3
7	**K**	R	C	**R**	K	SIPP	I	C	F	PD	OH	9.8	0.5
8	**K**	A	C	**R**	K	SIPP	R	C	H	PD	OH	1301	53
9	**K**	R	C	**R**	K	SIAP	R	C	H	PD	OH	2078	62
10	**K**	R	C	**K**	K	SIPP	R	C	H	PD	OH	3.8	0.2
11	**K**	R	C	**K**	K	SIPP	I	C	F		NH_2	0.49	0.04
12	**K**	R	C	**K**	K	SIPP	I	C			NH_2	0.71	0.04
13	**K**	R	C	**K**	K	SIPP	R	C			NH_2	4.7	0.1
14	**K**	R	C	**K**	K	SIPP	A	C	F		NH_2	3.7	0.2
15	**K**	R	A	**R**	K	SIPP	R	A	H	PD	OH	2373	87
16[a]	**K**	R	Aha	**K**	K	SIPP	I	Pra			NH_2	21.8	0.6
17[b]	**K**	R	Aha	**K**	K	SIPP	I	Pra			NH_2	5.0	0.4

[a] Compound **16** was cyclized by CuAAC. [b] Compound **17** was cyclized by RuAAC.

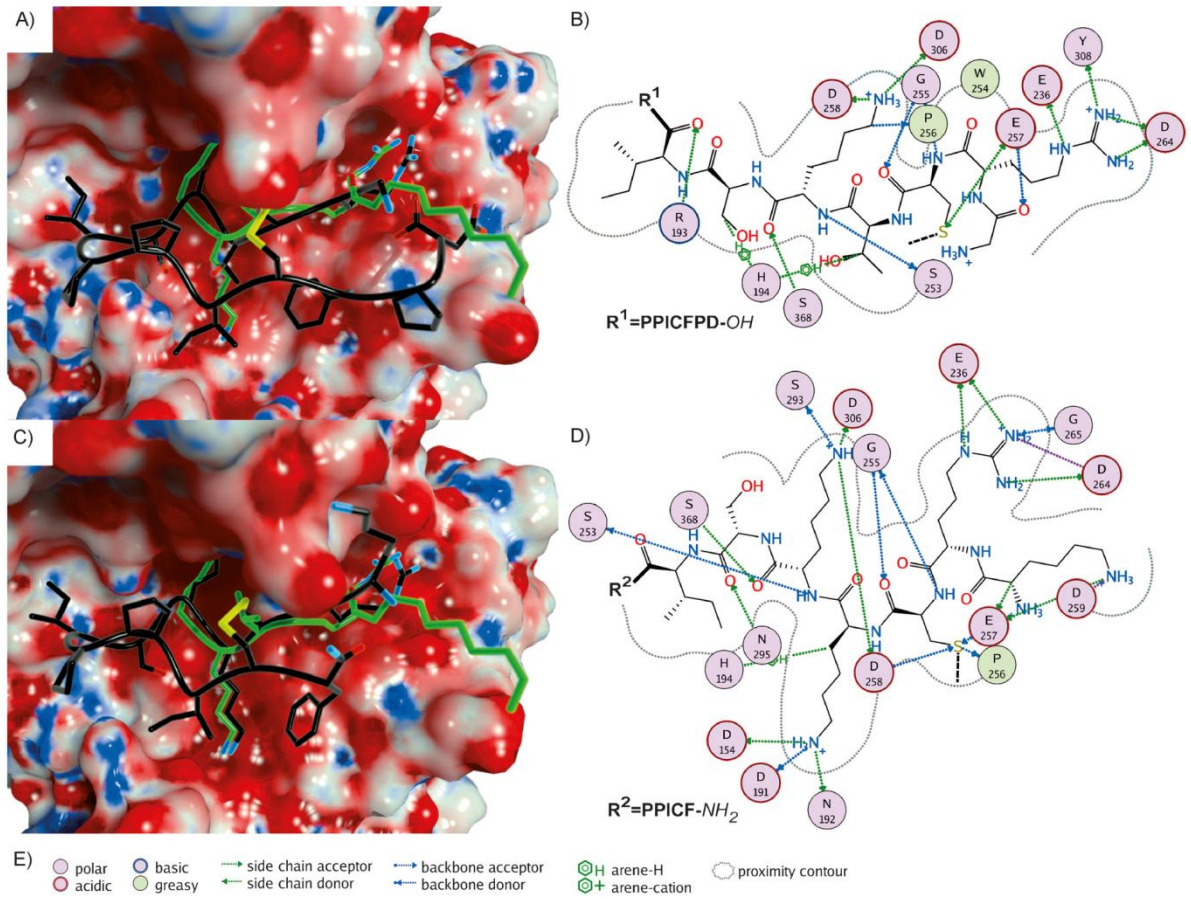


Figure 1. Models of furin in complex with A) and B) **1**, and C) and D) **11**. A) and C) Furin–ligand structures. Protein surfaces are colored according to the electrostatic surface potential (red: negative; blue: positive; calculated with the "Poisson Boltzmann Solver" algorithm). Ligands **1** and **11** shown in black with co-crystallized inhibitors in green (from PDB ID: 1P8J);[25] side-chain and terminal hetero atoms: blue = nitrogen, red = oxygen, yellow = sulfur (hydrogens omitted for clarity). B) and D) 2D depiction of main interactions of residues 1–7 of **1** and **11**. Proximal furin residues with no detected favorable interactions, water molecules, and ions are omitted for clarity. The disulfide bridge is indicated as a dashed line. E) Legend.

Finally, the role of the structure-stabilizing disulfide bridge was examined by replacing the cysteines with alanines (linear peptide **15**). This had dramatically reduced potency (K_i = 2.3 μM). Additionally, we replaced the disulfide bond with triazole bridges of different architectures (distinct substitution patterns). We previously demonstrated that SFTI-derived peptides bearing a 1,4-disubstituted 1,2,3-triazole as an ersatz disulfide bridge showed nearly no inhibitory activity towards trypsin, although their 1,5-disubstituted counterparts were good inhibitors.[24] In these redox-stable structural disulfide mimics, the activity of the peptide mimetic strongly depends on the steric resemblance to the natural peptidic inhibitor.[24] Thus we synthesized **16** and **17** with differently substituted triazoles. As described previously, on-resin copper(I)- or ruthenium(II)-catalyzed azide-alkyne cycloaddition (CuAAC or RuAAC) between L-azidohomoalanine (Aha) at position 3 and L-propargylglycine (Pra) at position 11 resulted in 1,4- (**16**) and 1,5-disubstituted (**17**) 1,2,3-triazolyl peptides, respectively. As in the trypsin assays, the RuAAC-assembled variant **17** appeared to be a more potent furin inhibitor than the CuAAC-synthesized compound **16** (K_i = 5 nM for **17**, 21.8 nM for **16**). In addition, both **16** and **17** demonstrated selectivity for furin, as indicated by drastically higher K_i values observed in inhibition assays with matriptase-1 and trypsin (Table 2). Although the divergence in potency between the two variants in the case of furin inhibition was not as dramatic as observed for trypsin,[24] **17** is a promising redox-stable candidate for future in vitro and in vivo experiments.

Finally, the selectivity of **11** and **12** towards furin was examined in assays with the serine protease matriptase-1 (K_i = 0.56 and 4.3 μM , respectively; Table 2). This is in agreement with our previous studies, which showed the importance of residue 12 for interaction with matriptase-1.[30] Hence, this position provides a handle to tune potency as well as selectivity of this versatile serine protease inhibitor scaffold.

In conclusion, we have rationally designed potent inhibitors of the therapeutically relevant subtilisin-like serine protease furin, based on a monocyclic variant of sunflower trypsin inhibitor. Although the bicyclic variant of the SFTI scaffold offers some special features, such as superior stability towards pro-

Table 2. Selectivity of the most promising furin binders.

	K_i [nM]		
	Matriptase	Trypsin	Furin
11	560 ± 36	> 10000	0.49
12	4528 ± 428	> 10000	0.71
16	9298 ± 308	> 10000	21.8
17	6681 ± 200	> 10000	5.0

teolysis and better activity against trypsin, we chose a peptide lacking head-to-tail cyclization, as its C-terminal residues do not really interact with the surface of furin, and thus providing the possibility to truncate the inhibitor. In contrast to targeting trypsin, this affords a decrease in molecular weight without sacrificing potency against the target protease. Rather, it appeared further improved, thus granting a significantly enhanced ligand binding efficiency. This strategy would not be possible with the bicyclic variant. In addition, the free amino terminus of the monocyclic scaffold provides the possibility to readily attach, for example, fluorescent labels[23] or other reporter moieties; this might be useful for furin-targeting biochemical tools in the future.

Although not as potent as 4-aminomethylbenzamidine-derived small molecules, these peptidic compounds could represent a significant contribution to the toolbox of PC inhibitors as they possess several advantageous features. Indeed, they can be easily conjugated with binding biomolecules, for example, antibodies, both chemically and upon recombinant production, in order to enhance half-life and assure targeted delivery in living organism. Moreover, triazolyl-containing inhibitors combine excellent potency and selectivity with enhanced redox and proteolytic stability[31] as a result of a non-natural scaffolding motif, and thus could become a viable tool to study and control PC-related processes.

Acknowledgements

This work was supported in part by the Deutsche Forschungsgemeinschaft SPP1623 through grant KO1390/10-1.

Keywords: Bowman-Birk inhibitor · furin · protein engineering · rational design · SFTI-1 · structure–activity relationships

[1] N. G. Seidah, R. Day, M. Marcinkiewicz, M. Chrétien, *Ann. N. Y. Acad. Sci.* **1998**, *839*, 9–24.
[2] C. Thacker, A. M. Rose, *BioEssays* **2000**, *22*, 545–553.
[3] A. J. Roebroek, J. A. Schalken, J. A. Leunissen, C. Onnekink, H. P. Bloemers, W. J. Van de Ven, *EMBO J.* **1986**, *5*, 2197–2202.
[4] W. M. van de Ven, J. Voorberg, R. Fontijn, H. Pannekoek, A. W. van den Ouweland, H. L. P. van Duijnhoven, A. M. Roebroek, R. Siezen, *Mol. Biol. Rep.* **1990**, *14*, 265–275.
[5] M. Chrétien, N. G. Seidah, A. Basak, M. Mbikay, *Expert Opin. Ther. Targets* **2008**, *12*, 1289–1300.
[6] E. D. Anderson, S. S. Molloy, F. Jean, H. Fei, S. Shimamura, G. Thomas, *J. Biol. Chem.* **2002**, *277*, 12879–12890.
[7] S. Tian, Q. Huang, Y. Fang, J. Wu, *Int. J. Mol. Sci.* **2011**, *12*, 1060–1065.
[8] N. G. Seidah, A. Prat, *Nat. Rev. Drug Discovery* **2012**, *11*, 367–383.
[9] A. J. M. Roebroek, L. Umans, I. G. L. Pauli, E. J. Robertson, F. van Leuven, W. J. M. Van de Ven, D. B. Constam, *Development* **1998**, *125*, 4863–4876.
[10] N. A. Taylor, W. J. M. van de Ven, J. W. M. Creemers, *FASEB J.* **2003**, *17*, 1215–1227.
[11] F. Couture, F. D'Anjou, R. Day, *Biomol. Concepts* **2011**, *2*, 421–438.
[12] G. Thomas, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 753–766.
[13] F. Jean, K. Stella, L. Thomas, G. Liu, Y. Xiang, A. J. Reason, G. Thomas, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 7293–7298.
[14] F. Couture, A. Kwiatkowska, Y. L. Dory, R. Day, *Expert Opin. Ther. Pat.* **2015**, *25*, 379–396.
[15] A. Cameron, J. Appel, R. A. Houghten, I. Lindberg, *J. Biol. Chem.* **2000**, *275*, 36741–36749.
[16] M. M. Kacprzak, J. R. Peinado, M. E. Than, J. Appel, S. Henrich, G. Lipkind, R. A. Houghten, W. Bode, I. Lindberg, *J. Biol. Chem.* **2004**, *279*, 36788–36794.
[17] S. Hallenberger, V. Bosch, H. Angliker, E. Shaw, H.-D. Klenk, W. Garten, *Nature* **1992**, *360*, 358–361.
[18] W. Garten, A. Stieneke, E. Shaw, P. Wikstrom, H.-D. Klenk, *Virology* **1989**, *172*, 25–31.
[19] G.-S. Jiao, L. Cregar, J. Wang, S. Z. Millis, C. Tang, S. O'Malley, A. T. Johnson, S. Sareth, J. Larson, G. Thomas, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19707–19712.
[20] G. L. Becker, F. Sielaff, M. E. Than, I. Lindberg, S. Routhier, R. Day, Y. Lu, W. Garten, T. Steinmetzer, *J. Med. Chem.* **2010**, *53*, 1067–1075.
[21] G. L. Becker, Y. Lu, K. Hardes, B. Strehlow, C. Levesque, I. Lindberg, K. Sandvig, U. Bakowsky, R. Day, W. Garten, T. Steinmetzer, *J. Biol. Chem.* **2012**, *287*, 21992–22003.
[22] S. O. Dahms, K. Hardes, G. L. Becker, T. Steinmetzer, H. Brandstetter, M. E. Than, *ACS Chem. Biol.* **2014**, *9*, 1113–1118.
[23] H. Fittler, O. Avrutina, M. Empting, H. Kolmar, *J. Pept. Sci.* **2014**, *20*, 415–420.
[24] M. Empting, O. Avrutina, R. Meusinger, S. Fabritz, M. Reinwarth, M. Biesalski, S. Voigt, G. Buntkowsky, H. Kolmar, *Angew. Chem. Int. Ed.* **2011**, *50*, 5207–5211; *Angew. Chem.* **2011**, *123*, 5313–5317.
[25] S. Henrich, A. Cameron, G. P. Bourenkov, R. Kiefersauer, R. Huber, I. Lindberg, W. Bode, M. E. Than, *Nat. Struct. Mol. Biol.* **2003**, *10*, 520–526.
[26] S. S. Molloy, P. A. Bresnahan, S. H. Leppla, K. R. Klimpel, G. Thomas, *J. Biol. Chem.* **1992**, *267*, 16396–16402.
[27] J. A. Walker, S. S. Molloy, G. Thomas, T. Sakaguchi, T. Yoshida, T. M. Chambers, Y. Kawaoka, *J. Virol.* **1994**, *68*, 1213–1218.
[28] I. Schechter, A. Berger, *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.
[29] M. Tischler, D. Nasu, M. Empting, S. Schmelz, D. W. Heinz, P. Rottmann, H. Kolmar, G. Buntkowsky, D. Tietze, O. Avrutina, *Angew. Chem. Int. Ed.* **2012**, *51*, 3708–3712; *Angew. Chem.* **2012**, *124*, 3768–3772.
[30] H. Fittler, O. Avrutina, B. Glotzbach, M. Empting, H. Kolmar, *Org. Biomol. Chem.* **2013**, *11*, 1848–1857.
[31] A. Gori, C.-I. A. Wang, P. J. Harvey, K. J. Rosengren, R. F. Bhola, M. L. Gelmi, R. Longhi, M. J. Christie, R. J. Lewis, P. F. Alewood, A. Brust, *Angew. Chem. Int. Ed.* **2015**, *54*, 1361–1364; *Angew. Chem.* **2015**, *127*, 1378–1381.

Manuscript received: August 31, 2015
Accepted article published: October 1, 2015
Final article published: October 19, 2015

3. Summary

This cumulative thesis relies on the studies comprising structure-based drug design towards novel, potent inhibitors of disease-related serine proteases on the scaffold of sunflower trypsin inhibitor-1 (SFTI-1) peptide.

Our initial study (Org. Biomol. Chem., **2012**, 10, 7753-7762) revealed the utility of SFTI-1 as a scaffold for computational design. Thus, the replacement of an important structural motif of SFTI[*1*,*14*], namely its disulfide bridge, against differently substituted 1,2,3-triazoles showed that inhibitory activity against a model protease trypsin was only preserved if the resulting architecture matched that of the parent peptide. Therefore, in the following investigation these *in-silico* modeled synthetic compounds were examined in the assays with the disease-related peptidase matriptase-1 and showed marginal affinity to this target. It was a surprising outcome taking into consideration that surface of matriptase-1 is negatively charged around the active site and SFTI-1[*1*,*14*] is charged positively. To explain this effect, the *in silico* calculated free energies for every inhibitor-enzyme complex were compared with those resulted from the *in vitro* data, and almost a perfect match was obtained. Therefore, the feasible explanation for the impaired binding could be an entropic penalty from the *C*-terminal loop region. Therefore, its truncation or replacement of certain amino acids could increase the binding potency towards matriptase-1.

The data obtained from the *in silico* study implied usage of the monocyclic version of SFTI-1 (SFTI-1[*1*,*14*]) as starting scaffold towards generation of more potent matriptase-1 inhibitors. From the crystal structure of the matriptase-1-SFTI-1 complex, three positions within the peptide were identified for substitutions. Thus, amino acids Ile7 and Ile10 were exchanged against non-natural azide-bearing amino acids and successive copper-catalyzed azide-alkyne cycloaddition (CuAAC) with different alkyne counterparts resulted in the library of 22 peptide mimetics. This 1,2,3-triazole-bearing compounds were examined in the inhibition assays against matriptase-1. Additionally, Phe12 was replaced by natural amino acids with less bulky side chains. Only the substitutions at position 10 and 12 led to improved, compared to the wildtype SFTI-1[*1*,*14*], affinity. The triazolyl amine at position 10 was replaced by positively charged canonical amino acids Lys, respectively, Arg. Surprisingly, both were more potent than the parent compound, with the arginine-bearing one being the most active inhibitor. The peptide possessing a combination of the two most beneficial replacements and a K_i of 11 nM (703 nM for the wildtype SFTI-1[*1*,*14*]) was named SFTI-1-derived matriptase inhibitor-1 (SDMI-1). This engineered SDMI-1 peptide contained exclusively canonical amino acids and was readily accessible by automated Fmoc-SPPS.

The next study was focused on the additional addressable site for the installation of tailor-made functionalities. This is needed to get access to oligomeric inhibitors *via* e.g. covalent grafting onto certain oligovalent scaffolds like an antibody or a C4b-binding protein (all seven α-helices of C4bp can be addressed). To this end, different positions were examined towards exchanges, but the decline of inhibitory activity against matriptase-1 was too significant. Obviously, attachment of novel functionalities could be achieved *via* the free *N*-terminus which is easily accessible by standard amide coupling. However, the elongation by a fluorophore group at the beginning of the peptide sequence resulted in a decrease of the inhibitory activity (K_i = 328 nM). On the other hand, the potency of the head-to-tail cyclized peptide was not impaired. Therefore, the decrease of binding capacity was presumably caused by a repulsion of the peptidase's and inhibitor's surfaces rather than by missing positive charge at the amino terminus. Therefore, an ε-Fmoc-protected lysine was incorporated at the first position and, following deprotection, its side-chain was used for the installation of desired functionalities. This new molecule, being with its K_i of 2.1 nM more potent as the precursor, was called SDMI-3. The coupling of different reporters, e.g. fluorophores or other motifs, showed only minor loss of potency against matriptase-1. Therefore, it was possible to conjugate SDMI-3 to different oligovalent biomolecular scaffolds, leading to tetrameric constructs upon coupling with the Fc part of an antibody or even heptamers in the case of the C4bp. These results are not included in this work.

The improved variants obtained in the previous studies were used as a starting point for the development of novel inhibitors of furin, a protease associated with Alzheimer`s disease, cancer, atherosclerosis and other pathologies. The wildtype SFTI-1[*1*,*14*] showed only moderate inhibition of furin (K_i = 35 μM) but SDMI-3 had a potency of 24.1 nM and became a good lead. The incorporation of the furin substrate sequence (Arg-X-Arg/Lys-Arg↓) did not result in a better binding. Interestingly, although in all furin inhibitors an Arg at the P1 position is fixed, the substitution against a Lys resulted in a much more potent compound. The replacement of neutral amino acids against positively charged ones showed improved binding with the negatively charged surface of the peptidase. Furthermore, the *in silico* model showed no pronounced interaction between the *C*-terminal region and the furin surface. Hence, truncated versions lacking amino acids 13-14 and 12-14 were synthesized. Both deletions had a prominent effect on the activity, leading to sub-nanomolar K_i's (0.49 and 0.71 nM, respectively), and the most active compound was named SFTI-derived furin inhibitor (SDFI).

The engineered peptides SDMI-3 and SDFI are valuable leads that can be used as scaffolds for in radio-labeling, liquid scintillation counting (LSC), single-photon emission computed tomography (SPECT) and positron emission tomography (PET).

4. Literature

[1] P. Cirino, K. Mayer, D. Umeno, in *Directed Evolution Library Creation, Vol. 231* (Eds.: F. Arnold, G. Georgiou), Humana Press, **2003**, pp. 3-9.

[2] L. Pritchard, D. Corne, D. Kell, J. Rowland, M. Winson, *Journal of Theoretical Biology* **2005**, *234*, 497-509.

[3] C. F. Barbas, J. D. Bain, D. M. Hoekstra, R. A. Lerner, *Proceedings of the National Academy of Sciences of the United States of America* **1992**, *89*, 4457-4461.

[4] C.-C. Yin, L.-L. Ren, L.-L. Zhu, X.-B. Wang, Z. Zhang, H.-L. Huang, X.-Y. Yan, *Journal of Biochemistry* **2008**, *144*, 591-598.

[5] J. Van den Brulle, Fischer, M., Langmann, T., Horn, G., Waldmann, T., Arnold, S., et al. (2008). *BioTechniques 45,* **2008**, *45*, 340-343.

[6] S. Chen, J. Qiu, C. Chen, C. Liu, Y. Liu, L. An, J. Jia, J. Tang, L. Wu, H. Hang, *Protein Cell* **2012**, *3*, 460-469.

[7] R. M. Schaaper, *Proceedings of the National Academy of Sciences of the United States of America* **1988**, *85*, 8126-8130.

[8] A. R. Timms, B. A. Bridges, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2002**, *499*, 97-101.

[9] H. Hu, J. Qian, J. Chu, Y. Wang, Y. Zhuang, S. Zhang, *Journal of Biotechnology* **2009**, *141*, 97-103.

[10] P. S. Daugherty, *Current Opinion in Structural Biology* **2007**, *17*, 474-480.

[11] J. Löfblom, *Biotechnology Journal* **2011**, *6*, 1115-1129.

[12] K. Hertveldt, T. Beliën, G. Volckaert, in *Bacteriophages, Vol. 502* (Eds.: M. J. Clokie, A. Kropinski), Humana Press, **2009**, pp. 321-339.

[13] G. P. Smith, *Science* **1985**, *228*, 1315-1317.

[14] R. W. Roberts, J. W. Szostak, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 12297-12302.

[15] C. A. Valencia, S. W. Cotten, B. Dong, R. Liu, *Biotechnology Progress* **2008**, *24*, 561-569.

[16] J. Hanes, L. Jermutus, S. Weber-Bornhauser, H. R. Bosshard, A. Plückthun, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 14130-14135.

[17] J. Hanes, A. Plückthun, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 4937-4942.

[18] R. H. Kimura, D. S. Jones, L. Jiang, Z. Miao, Z. Cheng, J. R. Cochran, *PLoS ONE* **2011**, *6*, e16112.

[19] K. D. W. E. T. Boder, *Nat Biotechnol.* **1997**, *15*, 553-557.

[20] A. C. Anderson, *Chemistry & Biology* **2003**, *10*, 787-797.

[21] *European Journal of Biochemistry* **1992**, *204*, 1-3.

[22] L. Polgár, *Cell. Mol. Life Sci.* **2005**, *62*, 2161-2172.

[23] N. D. Rawlings, A. J. Barrett, A. Bateman, *Nucleic Acids Research* **2012**, *40*, D343-D350.

[24] B. M. Dunn, *Chemical Reviews* **2002**, *102*, 4431-4458.

[25] M. Y. Kondo, D. N. Okamoto, J. A. N. Santos, M. A. Juliano, K. Oda, B. Pillai, M. N. G. James, L. Juliano, I. E. Gouvea, *Journal of Biological Chemistry* **2010**, *285*, 21437-21445.

[26] C. Tallant, A. Marrero, F. X. Gomis-Rüth, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2010**, *1803*, 20-28.

[27] N. Tajima, F. Kawai, S.-Y. Park, J. R. H. Tame, *Journal of Molecular Biology* **2010**, *402*, 645-656.

[28] C. Lopez-Otin, L. M. Matrisian, *Nat Rev Cancer* **2007**, *7*, 800-808.

[29] A. L. E. Seemüller, D. Stock, J. Löwe, R. Huber, W. Baumeister, *Science* **1995**, *286*, 579-582.

[30] E. Di Cera, *IUBMB life* **2009**, *61*, 510-515.

[31] Harold A. Chapman, a. Richard J. Riese, G.-P. Shi, *Annual Review of Physiology* **1997**, *59*, 63-88.

[32] Y. E. Shi, J. Torri, L. Yieh, A. Wellstein, M. E. Lippman, R. B. Dickson, *Cancer Research* **1993**, *53*, 1409-1415.
[33] J. Y. Kang, M. Dolled-Filhart, I. T. Ocal, B. Singh, C.-Y. Lin, R. B. Dickson, D. L. Rimm, R. L. Camp, *Cancer Research* **2003**, *63*, 1101-1105.
[34] A. Welman, D. Sproul, P. Mullen, M. Muir, A. R. Kinnaird, D. J. Harrison, D. Faratian, V. G. Brunton, M. C. Frame, *PLoS ONE* **2012**, *7*, e34182.
[35] H. Tanimoto, K. Shigemasa, X. Tian, L. Gu, J. B. Beard, T. Sawasaki, T. J. O'Brien, *Br J Cancer* **2004**, *92*, 278-283.
[36] I. Seitz, S. Hess, H. Schulz, R. Eckl, G. Busch, H. P. Montens, R. Brandl, S. Seidl, A. Schömig, I. Ott, *Arteriosclerosis, Thrombosis, and Vascular Biology* **2007**, *27*, 769-775.
[37] J. M. Milner, A. Patel, R. K. Davidson, T. E. Swingler, A. Desilets, D. A. Young, E. B. Kelso, S. T. Donell, T. E. Cawston, I. M. Clark, W. R. Ferrell, R. Plevin, J. C. Lockhart, R. Leduc, A. D. Rowan, *Arthritis & Rheumatism* **2010**, *62*, 1955-1966.
[38] M. L. L. Gao, N. Dong, Y. Jiang, C-Y. Lin, M. Huang, D. Wu and Q. Wu, *Leukemia* **2013**, *27*, 1191-1194.
[39] A. D. Santin, S. Cane, S. Bellone, E. Bignotti, M. Palmieri, L. E. De Las Casas, S. Anfossi, J. J. Roman, T. O'Brien, S. Pecorelli, *Cancer* **2003**, *98*, 1898-1904.
[40] K. List, R. Szabo, A. Molinolo, V. Sriuranpong, V. Redeye, T. Murdock, B. Burke, B. S. Nielsen, J. S. Gutkind, T. H. Bugge, *Genes & Development* **2005**, *19*, 1934-1950.
[41] M. Saleem, V. M. Adhami, W. Zhong, B. J. Longley, C.-Y. Lin, R. B. Dickson, S. Reagan-Shaw, D. F. Jarrard, H. Mukhtar, *Cancer Epidemiology Biomarkers & Prevention* **2006**, *15*, 217-227.
[42] B. S. K. Uhland, C. Arkona, M. Schuster, B. Petri, P. Steinmetzer, F. Mueller, A. Schweinitz, T. Steinmetzer, A. Van De Locht, *Int J Oncol* **2009**, *35*, 347-357.
[43] S.-Y. Choi, S. Bertram, I. Glowacka, Y. W. Park, S. Pöhlmann, *Trends in Molecular Medicine*, *15*, 303-312.
[44] G. Miller, K. List, *Cell Tissue Res* **2013**, *351*, 245-253.
[45] K. Uhland, *Cell. Mol. Life Sci.* **2006**, *63*, 2968-2978.
[46] M. D. Oberst, L.-Y. L. Chen, K.-I. Kiyomiya, C. A. Williams, M.-S. Lee, M. D. Johnson, R. B. Dickson, C.-Y. Lin, *American Journal of Physiology - Cell Physiology* **2005**, *289*, C462-C470.
[47] M. D. Oberst, B. Singh, M. Ozdemirli, R. B. Dickson, M. D. Johnson, C.-Y. Lin, *Journal of Histochemistry & Cytochemistry* **2003**, *51*, 1017-1025.
[48] I. Schechter, A. Berger, *Biochemical and Biophysical Research Communications* **1967**, *27*, 157-162.
[49] B. Zhao, C. Yuan, R. Li, D. Qu, M. Huang, J. C. K. Ngo, *Journal of Biological Chemistry* **2013**, *288*, 11155-11164.
[50] T. Takeuchi, J. L. Harris, W. Huang, K. W. Yan, S. R. Coughlin, C. S. Craik, *Journal of Biological Chemistry* **2000**, *275*, 26333-26342.
[51] S.-L. Lee, R. B. Dickson, C.-Y. Lin, *Journal of Biological Chemistry* **2000**, *275*, 36720-36725.
[52] I. J. Enyedy, S.-L. Lee, A. H. Kuo, R. B. Dickson, C.-Y. Lin, S. Wang, *Journal of Medicinal Chemistry* **2001**, *44*, 1349-1355.
[53] T. Steinmetzer, A. Schweinitz, A. Stürzebecher, D. Dönnecke, K. Uhland, O. Schuster, P. Steinmetzer, F. Müller, R. Friedrich, M. E. Than, W. Bode, J. Stürzebecher, *Journal of Medicinal Chemistry* **2006**, *49*, 4116-4126.
[54] A. Schweinitz, D. Dönnecke, A. Ludwig, P. Steinmetzer, A. Schulze, J. Kotthaus, S. Wein, B. Clement, T. Steinmetzer, *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 1960-1965.
[55] É. Colombo, A. Désilets, D. Duchêne, F. Chagnon, R. Najmanovich, R. Leduc, E. Marsault, *ACS Medicinal Chemistry Letters* **2012**, *3*, 530-534.
[56] B. Glotzbach, M. Reinwarth, N. Weber, S. Fabritz, M. Tomaszowski, H. Fittler, A. Christmann, O. Avrutina, H. Kolmar, *PLoS ONE* **2013**, *8*, e76956.
[57] F. M. Franco, D. E. Jones, P. K. W. Harris, Z. Han, S. A. Wildman, C. M. Jarvis, J. W. Janetka, *Bioorganic & Medicinal Chemistry* **2015**, *23*, 2328-2343.

[58] R. Goswami, G. Wohlfahrt, S. Mukherjee, C. Ghadiyaram, J. Nagaraj, L. K. Satyam, K. Subbarao, S. Gopinath, N. R. Krishnamurthy, H. S. Subramanya, M. Ramachandra, *Bioorganic & Medicinal Chemistry Letters* **2015**, *25*, 616-620.
[59] Z. Han, P. K. W. Harris, D. E. Jones, R. Chugani, T. Kim, M. Agarwal, W. Shen, S. A. Wildman, J. W. Janetka, *ACS Medicinal Chemistry Letters* **2014**, *5*, 1219-1224.
[60] R. Goswami, S. Mukherjee, C. Ghadiyaram, G. Wohlfahrt, R. K. Sistla, J. Nagaraj, L. K. Satyam, K. Subbarao, R. K. Palakurthy, S. Gopinath, N. R. Krishnamurthy, T. Ikonen, A. Moilanen, H. S. Subramanya, P. Kallio, M. Ramachandra, *Bioorganic & Medicinal Chemistry* **2014**, *22*, 3187-3203.
[61] S. J. de Veer, C. K. Wang, J. M. Harris, D. J. Craik, J. E. Swedberg, *Journal of Medicinal Chemistry* **2015**, *58*, 8257-8268.
[62] R. Goswami, S. Mukherjee, G. Wohlfahrt, C. Ghadiyaram, J. Nagaraj, B. R. Chandra, R. K. Sistla, L. K. Satyam, D. S. Samiulla, A. Moilanen, H. S. Subramanya, M. Ramachandra, *ACS Medicinal Chemistry Letters* **2013**, *4*, 1152-1157.
[63] N. A. TAYLOR, W. J. M. VAN DE VEN, J. W. M. CREEMERS, *The FASEB Journal* **2003**, *17*, 1215-1227.
[64] F. Couture, F. D'Anjou, R. Day, *Biomolecular concepts* **2011**, *2*, 421-438.
[65] H.-Y. Wang, D. H. S. Lee, M. R. D'Andrea, P. A. Peterson, R. P. Shank, A. B. Reitz, *Journal of Biological Chemistry* **2000**, *275*, 5626-5632.
[66] G. Thomas, *Nat Rev Mol Cell Biol* **2002**, *3*, 753-766.
[67] F. Couture, A. Kwiatkowska, Y. L. Dory, R. Day, *Expert Opinion on Therapeutic Patents* **2015**, *25*, 379-396.
[68] S. Henrich, A. Cameron, G. P. Bourenkov, R. Kiefersauer, R. Huber, I. Lindberg, W. Bode, M. E. Than, *Nat Struct Mol Biol* **2003**, *10*, 669-669.
[69] S. Hallenberger, V. Bosch, H. Angliker, E. Shaw, H.-D. Klenk, W. Garten, *Nature* **1992**, *360*, 358-361.
[70] A. S. W. Garten, E. Shaw, P. Wikstrom, HD. Klenk, *Virology* **1989**, *172*, 25-31.
[71] F. Jean, K. Stella, L. Thomas, G. Liu, Y. Xiang, A. J. Reason, G. Thomas, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 7293-7298.
[72] A. Cameron, J. Appel, R. A. Houghten, I. Lindberg, *Journal of Biological Chemistry* **2000**, *275*, 36741-36749.
[73] M. M. Kacprzak, J. R. Peinado, M. E. Than, J. Appel, S. Henrich, G. Lipkind, R. A. Houghten, W. Bode, I. Lindberg, *Journal of Biological Chemistry* **2004**, *279*, 36788-36794.
[74] G.-S. Jiao, L. Cregar, J. Wang, S. Z. Millis, C. Tang, S. O'Malley, A. T. Johnson, S. Sareth, J. Larson, G. Thomas, *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 19707-19712.
[75] G. L. Becker, F. Sielaff, M. E. Than, I. Lindberg, S. Routhier, R. Day, Y. Lu, W. Garten, T. Steinmetzer, *Journal of Medicinal Chemistry* **2010**, *53*, 1067-1075.
[76] G. L. Becker, Y. Lu, K. Hardes, B. Strehlow, C. Levesque, I. Lindberg, K. Sandvig, U. Bakowsky, R. Day, W. Garten, T. Steinmetzer, *The Journal of Biological Chemistry* **2012**, *287*, 21992-22003.
[77] R. Jefferis, *Expert Opinion on Biological Therapy* **2007**, *7*, 1401-1413.
[78] S. Aggarwal, *Nat Biotech* **2014**, *32*, 32-39.
[79] M. L. Colgrave, D. J. Craik, *Biochemistry* **2004**, *43*, 5965-5975.
[80] N. V. C. Shelley E Ackerman , Jamie M Bergen , Jennifer R Cochran, *Expert Review of Proteomics* **2014**, *11*.
[81] A. H. L. Chiche, J. Gelly, J. Gracy, P. Chau, P. Ha, J. Hernandez, D. Le-Nguyen, *Curr Protein Pept Sci* **2004**, *5*, 341-349.
[82] A. Heitz, O. Avrutina, D. Le-Nguyen, U. Diederichsen, J.-F. Hernandez, J. Gracy, H. Kolmar, L. Chiche, *BMC Structural Biology* **2008**, *8*, 54.
[83] D. J. Craik, N. L. Daly, T. Bond, C. Waine, *Journal of Molecular Biology* **1999**, *294*, 1327-1336.
[84] H. Kolmar, *Current Pharmaceutical Design* **2011**, *17*, 4329-4336.

[85] A. Heitz, J.-F. Hernandez, J. Gagnon, T. T. Hong, T. T. C. Pham, T. M. Nguyen, D. Le-Nguyen, L. Chiche, *Biochemistry* **2001**, *40*, 7973-7983.
[86] R. Eliasen, N. L. Daly, B. S. Wulff, T. L. Andresen, K. W. Conde-Frieboes, D. J. Craik, *Journal of Biological Chemistry* **2012**, *287*, 40493-40501.
[87] P. J. Jackson, J. C. McNulty, Y.-K. Yang, D. A. Thompson, B. Chai, I. Gantz, G. S. Barsh, G. L. Millhauser, *Biochemistry* **2002**, *41*, 7565-7572.
[88] T. S. M. Werle, H. Huang, A. Wentzel, H. Kolmar, A. Bernkop-Schnürch, *Journal of Drug Targeting* **2006**, *14*, 137-146.
[89] K. P. Greenwood, N. L. Daly, D. L. Brown, J. L. Stow, D. J. Craik, *The International Journal of Biochemistry & Cell Biology* **2007**, *39*, 2252-2264.
[90] S. Reiss, M. Sieber, V. Oberle, A. Wentzel, P. Spangenberg, R. Claus, H. Kolmar, W. Lösche, *Platelets* **2006**, *17*, 153-157.
[91] A. Christmann, K. Walter, A. Wentzel, R. Krätzner, H. Kolmar, *Protein Engineering* **1999**, *12*, 797-806.
[92] A. P. Silverman, A. M. Levin, J. L. Lahti, J. R. Cochran, *Journal of Molecular Biology* **2009**, *385*, 1064-1075.
[93] J. A. Getz, O. Cheneval, D. J. Craik, P. S. Daugherty, *ACS Chemical Biology* **2013**, *8*, 1147-1154.
[94] N. L. Daly, Y.-K. Chen, F. M. Foley, P. S. Bansal, R. Bharathi, R. J. Clark, C. P. Sommerhoff, D. J. Craik, *Journal of Biological Chemistry* **2006**, *281*, 23668-23675.
[95] A. B. E. Brauer, G. J. Domingo, R. M. Cooke, S. J. Matthews, R. J. Leatherbarrow, *Biochemistry* **2002**, *41*, 10608-10615.
[96] J. R. Costa, S. N. Yaliraki, *The Journal of Physical Chemistry B* **2006**, *110*, 18981-18988.
[97] M. Tischler, D. Nasu, M. Empting, S. Schmelz, D. W. Heinz, P. Rottmann, H. Kolmar, G. Buntkowsky, D. Tietze, O. Avrutina, *Angewandte Chemie International Edition* **2012**, *51*, 3708-3712.
[98] K. Hilpert, G. Hansen, H. Wessner, R. Volkmer-Engert, W. Höhne, *Journal of Biochemistry* **2005**, *138*, 383-390.
[99] S. Luckett, R. S. Garcia, J. J. Barker, A. V. Konarev, P. R. Shewry, A. R. Clarke, R. L. Brady, *Journal of Molecular Biology* **1999**, *290*, 525-533.
[100] D. Debowski, R. Łukajtis, M. Filipowicz, P. Strzelecka, M. Wysocka, A. Łęgowska, A. Lesner, K. Rolka, *Peptide Science* **2013**, *100*, 154-159.
[101] H. Fittler, A. Depp, O. Avrutina, S. O. Dahms, M. E. Than, M. Empting, H. Kolmar, *Chembiochem* **2015**, *16*, 2441-2444.
[102] H. Fittler, O. Avrutina, B. Glotzbach, M. Empting, H. Kolmar, *Organic & Biomolecular Chemistry* **2013**, *11*, 1848-1857.
[103] A. Łęgowska, E. Bulak, M. Wysocka, A. Jaśkiewicz, A. Lesner, D. Dębowski, K. Rolka, *Bioorganic & Medicinal Chemistry* **2008**, *16*, 5644-5652.
[104] P. Quimbar, U. Malik, C. P. Sommerhoff, Q. Kaas, L. Y. Chan, Y.-H. Huang, M. Grundhuber, K. Dunse, D. J. Craik, M. A. Anderson, N. L. Daly, *Journal of Biological Chemistry* **2013**, *288*, 13885-13896.
[105] M. L. J. Korsinczky, H. J. Schirra, K. J. Rosengren, J. West, B. A. Condie, L. Otvos, M. A. Anderson, D. J. Craik, *Journal of Molecular Biology* **2001**, *311*, 579-591.
[106] Y.-Q. Long, S.-L. Lee, C.-Y. Lin, I. J. Enyedy, S. Wang, P. Li, R. B. Dickson, P. P. Roller, *Bioorganic & Medicinal Chemistry Letters* **2001**, *11*, 2515-2519.
[107] J. D. McBride, H. N. M. Freeman, R. J. Leatherbarrow, *European Journal of Biochemistry* **1999**, *266*, 403-412.
[108] J. Austin, R. Kimura, Y.-H. Woo, J. Camarero, *Amino Acids* **2010**, *38*, 1313-1322.
[109] M. Empting, O. Avrutina, R. Meusinger, S. Fabritz, M. Reinwarth, M. Biesalski, S. Voigt, G. Buntkowsky, H. Kolmar, *Angewandte Chemie International Edition* **2011**, *50*, 5207-5211.
[110] E. Zabłotna, K. Kaźmierczak, A. Jaśkiewicz, M. Stawikowski, G. Kupryszewski, K. Rolka, *Biochemical and Biophysical Research Communications* **2002**, *292*, 855-859.
[111] X. Guo, J. Shi, Z. Tang, D. Cui, Y. Zhang, *Chemical Biology & Drug Design* **2006**, *68*, 341-344.

[112] P. Li, S. Jiang, S.-L. Lee, C. Y. Lin, M. D. Johnson, R. B. Dickson, C. J. Michejda, P. P. Roller, *Journal of Medicinal Chemistry* **2007**, *50*, 5976-5983.
[113] M. Stawikowski, R. Stawikowska, A. Jaśkiewicz, E. Zabłotna, K. Rolka, *Chembiochem* **2005**, *6*, 1057-1061.
[114] H. Fittler, O. Avrutina, M. Empting, H. Kolmar, *Journal of Peptide Science* **2014**, *20*, 415-420.
[115] R. García Boy, W. Mier, E. Nothelfer, A. Altmann, M. Eisenhut, H. Kolmar, M. Tomaszowski, S. Krämer, U. Haberkorn, *Mol Imaging Biol* **2010**, *12*, 377-385.
[116] A. Lesner, N. Karna, M. Psurski, A. Łęgowska, M. Wysocka, K. Guzow, A. Sieradzan, M. Sieńczyk, P. Trzonkowski, M. Pikuła, M. Zieliński, P. Kosikowska, R. Łukajtis, M. Łęgowska, D. Dębowski, W. Wiczk, K. Rolka, *Peptide Science* **2014**, *102*, 124-135.
[117] T. Uhlig, T. Kyprianou, F. G. Martinelli, C. A. Oppici, D. Heiligers, D. Hills, X. R. Calvo, P. Verhaert, *EuPA Open Proteomics* **2014**, *4*, 58-69.
[118] K. Fosgerau, T. Hoffmann, *Drug Discovery Today* **2015**, *20*, 122-128.
[119] D. Wu, Y. Gao, Y. Qi, L. Chen, Y. Ma, Y. Li, *Cancer Letters* **2014**, *351*, 13-22.
[120] S. Chernysh, I. Kozuharova, *International Immunopharmacology* **2013**, *17*, 1090-1093.
[121] S. J. Furlong, N. D. Ridgway, D. W. Hoskin, *International Journal of Oncology* **2008**, *32*, 537-544.
[122] J. Kang, G. Zhao, T. Lin, S. Tang, G. Xu, S. Hu, Q. Bi, C. Guo, L. Sun, S. Han, Q. Xu, Y. Nie, B. Wang, S. Liang, J. Ding, K. Wu, *Cancer Letters* **2013**, *339*, 247-259.
[123] B. Du, H. Han, Z. Wang, L. Kuang, L. Wang, L. Yu, M. Wu, Z. Zhou, M. Qian, *Molecular Cancer Research* **2010**, *8*, 135-144.
[124] T. Li, W. Luo, D. He, R. Wang, Y. Huang, X. Zeng, W. Wang, X. Chen, S. Gao, Y. Yu, X. Li, X. Wu, *Cancer Letters* **2013**, *339*, 226-236.
[125] E. Lee, J. E. Koskimaki, N. B. Pandey, A. S. Popel, *Neoplasia (United States)* **2013**, *15*, 112-124.
[126] A. Łęgowska, D. Dębowski, A. Lesner, M. Wysocka, K. Rolka, *Bioorganic & Medicinal Chemistry* **2009**, *17*, 3302-3307.
[127] A. Łęgowska, D. Dębowski, A. Lesner, M. Wysocka, K. Rolka, *Mol Divers* **2010**, *14*, 51-58.
[128] J. E. Swedberg, L. V. Nigon, J. C. Reid, S. J. de Veer, C. M. Walpole, C. R. Stephens, T. P. Walsh, T. K. Takayama, J. D. Hooper, J. A. Clements, A. M. Buckle, J. M. Harris, *Chemistry & Biology* **2009**, *16*, 633-643.
[129] E. Zabłotna, A. Jaśkiewicz, A. Łęgowska, H. Miecznikowska, A. Lesner, K. Rolka, *Journal of Peptide Science* **2007**, *13*, 749-755.
[130] J. D. McBride, E. M. Watson, A. B. E. Brauer, A. M. Jaulent, R. J. Leatherbarrow, *Peptide Science* **2002**, *66*, 79-92.
[131] H. J. V. Pereira, M. C. O. Salgado, E. B. Oliveira, *Journal of Chromatography B* **2009**, *877*, 2039-2044.
[132] J. E. Swedberg, S. J. de Veer, K. C. Sit, C. F. Reboul, A. M. Buckle, J. M. Harris, *PLoS ONE* **2011**, *6*, e19302.
[133] Ł. Rafał, Ł. Anna, W. Magdalena, D. Dawid, L. Adam, R. Krzysztof, *Journal of Peptide Science* **2011**, *17*, 281-287.
[134] A. Łęgowska, A. Lesner, E. Bulak, A. Jaśkiewicz, A. Sieradzan, M. Cydzik, P. Stefanowicz, Z. Szewczuk, K. Rolka, *FEBS Journal* **2010**, *277*, 2351-2359.
[135] C. Jendrny, A. G. Beck-Sickinger, *Chembiochem* **2015**, DOI: 10.1002/cbic.201500539.

5. Abbreviations

2-Cl-Trt	2-Chlorotrityl
Å	Ångström
aa	Amino acid
Ac	Acetate
Aha	Azidohomoalanine
Ala	Alanine
aq.	Aqueous
Arg	Arginine
Asp	Aspartic acid
BBI	Bowman-Birk inhibitor
Boc	*tert*-Butyloxycarbonyl
°C	Degree Celsius
calc.	Calculated
cp	Cylcopentadiene
CH_3CN	Acetonitrile
CuAAC	Copper(I)-catalyzed azide-alkyne cycloaddition
d	Day
Da	Dalton
DCM	Dichloromethane
DIEA	Ethyl diisopropyl amine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
[E]	Enzyme concentration
EI	Electron ionization
ELISA	Enzyme-linked immunosorbent assay
eq.	Equivalent
ESI	Electrospray ionization
et al.	*et alii, et aliae, et alia*
Fc	Fragment, crystallizable (antibody)
Fig.	Figure
FITC	Fluorescein isothiocyanate
Fmoc	9-Fluorenylmethyloxycarbonyl
G	Free energy
g	Gram
Glu	Glutamic acid
Gly	Glycine
h	Hour
H_2O	Water
HAI-1	Hepatocyte growth factor activator inhibitor-1
HBTU	*O*-(*1H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluroniumhexafluorophosphate
His	Histidine
HOAt	7-Aza-1-hydroxybenzotriazol
HOBt	1-Hydroxybenzotriazol
HPLC	High-performance liquid chromatography
[I]	Inhibitor concentration
Ig	Immunoglobulin
IC_{50}	Mean inhibitory concentration
Ile	Isoleucine
IR	Infrared
k	Kilo

K	Kelvin
K_i	Inhibition constant (substrate-independent)
K_i^{app}	Apparent inhibition constant (substrate-dependent)
K_m	Michaelis-Menten constant
L	Liter
Leu	Leucine
Lys	Lysine
λ	Wavelength
M	Molar concentration (mol/L)
m	Milli (10^{-3})
Me	Methyl
MeOH	Methanol
min	Minute
MS	Mass spectrometry
MTBE	*tert*-butyl-methyl ether
MW	Molecular weight
μ	Micro (10^{-6})
m/z	Mass-to-charge ratio
n	Nano (10^{-9})
$(NH_4)_2CO_3$	Ammonium carbonate
NMP	N-methyl-2-pyrrolidone
p	Pico (10^{-12})
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data base
PET	Positron emission tomography
Phe	Phenylalanine
PG	Protecting group
pH	$-\lg[H]^+$
ppm	Parts per million
Pra	Propargylglycine
Pro	Proline
PyBOP	Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
R	Residue
RMSD	Root-mean-square deviation
RP	Reversed-phase
R_t	Retention time
RuAAC	Ruthenium(II)-catalyzed azide-alkyne cycloaddition
SAR	Structure-activity relationship
SDFI	SFTI-1 derived furin inhibitor
SDMI-1	SFTI-1 derived matriptase inhibitor-1
SDMI-2	SFTI-1 derived matriptase inhibitor-2
SDMI-3	SFTI-1 derived matriptase inhibitor-3
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
Ser	Serine
SFTI-1	Sunflower trypsin inhibitor-1
SPPS	Solid phase peptide synthesis
T	Temperature
Tab	Table
TFA	Trifluoroacetic acid
TES	Triethylsilane
Trt	Trityl
Trp	Tryptophan

TTSP	Type II transmembrane serine protease
Tyr	Tyrosine
UV/Vis	Ultraviolet-visible
V	Volume
v	Initial velocity
Val	Valine

6. Supporting Information

This part contains the supporting information of all studies introduced in section 2. It provides additional experimental data and all information of the analytical methods used for the characterization of all synthesized compounds.

6.1 Supporting Information for:
Olga Avrutina, **Heiko Fittler**, Bernhard Glotzbach, Harald Kolmar, Martin Empting, Between two worlds: A comparative study on *in vitro* and *in silico* inhibition of trypsin and matriptase by redox-stable SFTI-1 variants at near physiological pH. Org Biomol Chem. **2012**, 10, 7753-7762.

6.2 Supporting Information for:
Heiko Fittler, Olga Avrutina, Bernhard Glotzbach, Martin Empting, Harald Kolmar, Combinatorial tuning of peptidic drug candidates: High-affinity matriptase inhibitors through incremental structure-guided optimization, *Org Biomol Chem.* **2013**, *11*, 1848-1857.

6.3 Supporting Information for:
Heiko Fittler, Olga Avrutina, Martin Empting, Harald Kolmar, Potent inhibitors of human matriptase-1 based on the scaffold of sunflower trypsin inhibitor, *J Pept Sci.* **2014**, *20*, 415-420.

6.4 Supporting Information for:
Heiko Fittler, Alexander Depp, Olga Avrutina, Sven Dahms, Manuel Than, Martin Empting, Harald Kolmar, Engineering a constrained peptidic scaffold towards potent and selective furin inhibitors, *ChemBioChem*, **2015**, *16*, 2441-2444.

6.1. Supporting Information for chapter 2.1

--Supporting Information--

Between Two Worlds: a Comparative Study on *In Vitro* and *In Silico* Inhibition of Trypsin and Matriptase by Redox-Stable SFTI-1 Variants at Near Physiological pH

Olga Avrutina,[a] Heiko Fittler,[a] Bernhard Glotzbach,[a] Harald Kolmar,[a] and Martin Empting[a,*]

[a] *Clemens-Schöpf Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Petersenstr. 22, 64287 Darmstadt, Germany.*

Table of Contents

Experimental Procedures 2

In Silico Methods 4

Supporting Figures 5

Supporting Tables 8

Experimental Procedures

Instrumentation

Analytical HPLC was conducted using a Varian 920-LC system equipped with a Phenomenex Hypersil 5u BDS C18 LC column (150 x 4.6 mm, 5 µm, 130 Å). Semi-preparative RP-HPLC was performed on a Varian modular system comprising a PrepStar 218 Solvent Delivery Module, a ProStar 410 HPLC AutoSampler and a ProStar 325 Dual Wavelength UV-Vis HPLC Detector using a YMC J'sphere ODS-H80 C-18 LC column (250 x 20 mm, 4 µm, 8 nm). The eluent system for analytical and semi-preparative HPLC consisted of eluent A (0.1% aq. TFA) and eluent B (90 % aq. acetonitrile containing 0.1% TFA).

ESI mass spectra were recorded using a Shimadzu LCMS-2020 equipped with a Phenomenex Jupiter 5u C4 LC column (50 x 1 mm, 5 µm, 300 Å). The eluent system consisted of eluent A (0.1% aq. formic acid, LC-MS grade) and eluent B (acetonitrile containing 0.1% formic acid, LC-MS grade).

Fast protein liquid chromatography (FPLC) was conducted using an ÄKTApurifier (Amersham Pharmacia Biotech).

General Fmoc-SPPS Procedures

Peptides were synthesized on a Liberty 12-channel automated peptide synthesizer coupled with a Discover SPS microwave peptide synthesizer platform (CEM) using the Fmoc strategy. Amino acids were attached by double or triple coupling employing 4 eq of the corresponding amino acid, 4 eq of 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronoium hexafluorphosphate (HBTU) and 8 eq of *N*,*N*-diisopropylethylamine (DIEA), or in case of cysteine 3-4 eq of 2,4,6-trimethylpyridine (collidine). Arginine and cysteine were coupled using a two-step microwave program: 1. RT, 0 W, 25 min; 2. 75 °C, 25 W, 0.5 min (Arg) and 1. RT, 0 W, 2 min; 2. 50 °C, 25 W, 4 min (Cys), respectively. All other amino acids were coupled using a standard microwave program: 75 °C, 21 W, 5 min.

Fmoc deprotection was achieved in two steps by reaction with 20% piperidine in DMF at 75 °C, 42 W for 0.5 min (initial deprotection) followed by a second deprotection step with 20% piperidine in DMF at 75 °C, 42 W for 3 min.

Compounds 2 and 4–6

Synthesis and characterization of compounds **2** and **4**–**6** have already been described in detail in reference 13 from the main text.

SFTI-1 (1)

The linear precursor of **1** was synthesized on chlorotrityl resin preloaded with Fmoc-Gly (0.59 mmol/g) at 0.25 mmol scale according to the automated Fmoc-SPPS protocol described above. The peptide was cleaved from the resin under conservation of side chain protection using 5 mL of a mixture of acetic acid, DCM, and methanol (50:40:10, v/v/v) for 2 h at ambient temperature. The solvents were evaporated. To the resulting yellow oil *n*-hexan was added and then evaporated. This step was repeated three times. The residue was dissolved in 20 mL H_2O/CH_3CN (1:1, v/v) and lyophilized. 30 mg (0.025 mmol) of the linear peptide were dissolved in 30 mL dry DMF, and 4.4 mg HOBT (5 eq), 17 mg (5 eq) benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) as well as 11.3 µL (10 eq) DIEA were added for backbone macrocyclization. After 16 h of reaction, additional portions of HOBT (4.4 mg, 5 eq), PyBOP (17 mg, 5 eq), and DIEA (11.3 µL, 10 eq) were added and the reaction mixture was stirred overnight at ambient temperature. The solvent was evaporated and the protecting groups were removed by acidolytic cleavage using TFA/H_2O/anisole/triethylsilane (TES) (47:1:1:1, v/v/v/v) and a small amount of dithiothreitol (DTT). The reaction mixture was shaken for 3 h at ambient temperature with subsequent by precipitation and washing (3×) with 30 mL methyl *tert*-butyl ether (MTBE) to yield the crude monocyclic peptide. Oxidative disulfide formation was conducted in 100 mM $(NH_4)_2CO_3$ aq (pH = 8.6) at 1 mg peptide/mL dilution. After complete conversion, the solvent was removed *in vacuo* to yield the crude peptide. Chromatographic isolation by RP-HPLC yielded 4.5 mg of pure **1** (11.9 %).

RP-HPLC: Rt = 15.5 min (18 % acetonitrile over 2 min followed by 18→40.5 % acetonitrile in 0.1 % TFA over 20 min at flow rate 1 mL/min). ESI-MS: m/z: $[M+H]^+$ obsd. = 1514.6 (calc = 1513.7), $[M+2H]^{2+}$ obsd. = 757.8 (calc = 757.4), $[M-H]^-$ obsd. = 1511.8 (calc = 1511.7).

[Ala^3(&^{1}),Ala^{11}(&^{2})]SFTI-1[1,14][(&^{1}-1,5-[1,2,3]triazolyl-&^{2})] (3)

The amino acid sequence Fmoc-Aza-Thr(tBu)-Lys(Boc)-Ser(tBu)-Ile-Pro-Pro-Ile-Pra-Phe-Pro-Asp(tBu) was assembled on an AmphiSpheres 40 HMP resin (0.4 mmol/g, Varian/Agilent) at 0.125 mmol scale using the automated microwave-assisted Fmoc-SPPS procedure described above. Loading of the resin with Fmoc-Asp(tBu)-OH was conducted by triple coupling 2 eq AA, 2 eq 2-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronoium hexafluorphosphate (HATU), 4 eq DIEA and two-step microwave program (1. 60 °C, 30 W, 45 min, 2. 75 °C, 20 W, 5 min). Fmoc-Aza-OH and Fmoc-Pra-OH were attached using double coupling of 2 eq AA, 2 eq HATU, 4 eq DIEA, and two-step microwave program (1. 60 °C, 30 W, 45 min, 2. 75 °C, 20 W, 5 min). On-support ruthenium(II)-catalyzed macrocyclization of linear resin-bound precursor was conducted as previously reported (reference 12 from the main text). *N*-terminal sequence Gly-Arg(Pbf) was assembled using double coupling for each amino acid (4 eq aa, 3.9 eq HBTU, 8 eq DIEA) and microwave

irradiation (50 °C, 30 W, 30 min). The peptide resin was dried and subjected to acidolytic cleavage using TFA/H_2O/anisole/TES (47:1:1:1, v/v/v/v). Ether precipitation, washing, and subsequent purification *via* semi-preparative HPLC yielded 2.1 mg macrocyclic peptide **3** (1.37 µmol, 1.1 % according to the initial loading of the resin).
RP-HPLC: Rt = 16.2 min (18 % acetonitrile over 2 min followed by 18→40.5 % acetonitrile in 0.1 % TFA over 20 min at flow rate 1 mL/min). ESI-MS (m/z) $[M+2H]^{2+}$ obsd. = 768.56 (calc = 767.9), $[M+3H]^{3+}$ obsd. = 512.85 (calc = 512.8), $[M-H]^-$ obsd. = 1533.29 (calc = 1532.8). IR (cm^{-1}) 3424, 2924, 1652, 1538, 1451, 1203, 1132.

Inhibition Assays

Kinetic curves were recorded by monitoring the absorption of the corresponding samples in 96-well plates (NUNC, flat bottom, clear) at 405 nm in intervals of 60 sec over 30 min at RT using the Tecan GENios microplate reader. All experiments were performed in triplicate. Trypsin from bovine pancreas (Sigma) or matriptase were standardized by active-site titration with *p*-nitrophenyl-*p*'-guanidinobenzoate (NPGB) in phosphate buffered saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10.0 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4).
The normalized residual proteolytic activity v/v_0 of trypsin against the chromogenic substrate Boc-QAR-*p*NA (250 µM, Bachem) at different concentrations of bicyclic and monocyclic SFTI-1 analogues **1-6** [I] was determined for ~ 0.5 nM (trypsin) or 0.9 nM (matriptase) active enzyme ([E]) in aqueous buffer (50 mM Tris/HCl, 150 mM NaCl, 0.01% Triton X-100, 0.01% sodium azide, pH 7.6 or 8.5). The apparent inhibition constants (K_i^{app}) were calculated by fitting the Morrison equation for tight binding inhibitors (1) onto the resulting kinetic data with the Marquardt-Levenberg algorithm of SigmaPlot 11.

Determination of Michaelis-Menten Constant (K_M) for Boc-QAR-pNA against Matriptase

The initial reaction rate (v_i) of the proteolytic degradation of Boc-QAR-*p*NA (Bachem) by matriptase (1 nM) was determined for a series of concentrations ($[S]_i$) of the chromogenic substrate (1000-75 µM for pH 7.6 and 1000-50 µM for pH 8.5). The Michaelis-Menten constant (K_M) was calculated *via* Lineweaver-Burk plot (reciprocal initial reaction rate 1/vi versus the reciprocal substrate concentration $1/[S]_i$) and linear regression of the resulting data. The experiment was performed in triplicate yielding K_M as 236.8±56.1 µM (pH 7.6) and 66.6±16.0 µM (pH 8.5) (arithmetic mean, standard deviation given as error).

Transformation of Equation (2) into Equation (3)

$$\frac{v}{v_o} = \frac{([E]-[I]-K_i^{app})+\sqrt{([I]+K_i^{app}-[E])^2+4K_i^{app}[E]}}{2[E]} \quad (2)$$

$$\frac{v}{v_o} = \frac{2[E]-2[E]+[E]-[I]-K_i^{app}+\sqrt{([I]+K_i^{app}-[E])^2+4K_i^{app}[E]}}{2[E]} \quad (8)$$

$$\frac{v}{v_o} = \frac{2[E]}{2[E]} + \frac{-[E]-[I]-K_i^{app}+\sqrt{([I]+K_i^{app}-[E])^2+4K_i^{app}[E]}}{2[E]} \quad (9)$$

$$\frac{v}{v_o} = 1 - \frac{([E]+[I]+K_i^{app})-\sqrt{([I]+K_i^{app}-[E])^2+4K_i^{app}[E]}}{2[E]} \quad (10)$$

$$\frac{v}{v_o} = 1 - \frac{([E]+[I]+K_i^{app})-\sqrt{[I]^2+(K_i^{app})^2+[E]^2+2K_i^{app}[I]-2[E][I]-2K_i^{app}[E]+4K_i^{app}[E]}}{2[E]} \quad (11)$$

$$\frac{v}{v_o} = 1 - \frac{([E]+[I]+K_i^{app})-\sqrt{[I]^2+[E]^2+(K_i^{app})^2+2K_i^{app}[E]+2K_i^{app}[I]+2[E][I]-2[E][I]-2[E][I]}}{2[E]} \quad (12)$$

trinomial theorem

$$\frac{v}{v_o} = 1 - \frac{([E]+[I]+K_i^{app})-\sqrt{([E]+[I]+K_i^{app})^2-4[E][I]}}{2[E]} \quad (3) \quad \blacksquare$$

Propagation of Errors for $\Delta_B G^{exp}$

$\Delta_B G^{exp}$ for trypsin and matriptase complexes of compounds **1**–**6** were calculated from *in vitro* K_i using equation (7). The error of $\Delta_B G^{exp}$ ($\Delta\Delta_B G^{exp}$) was calculated by propagation of errors of K_i (ΔK_i) as follows:

$$(\Delta\Delta_B G^{exp})^2 = \left|\frac{\partial \Delta_B G^{exp}}{\partial K_i}\right|^2 (\Delta K_i)^2 \quad (13)$$

$$\Delta\Delta_B G^{exp} = \sqrt{\left|\frac{\partial \Delta_B G^{exp}}{\partial K_i}\right|^2 (\Delta K_i)^2} \qquad (14)$$

Finally, differentiation of $\Delta_B G^{exp}$ with respect to K_i yields equation (15).

$$\Delta\Delta_B G^{exp} = \sqrt{\left|\frac{-RT}{K_i}\right|^2 (\Delta K_i)^2} \qquad (15)$$

In Silico Methods

Instrumentation

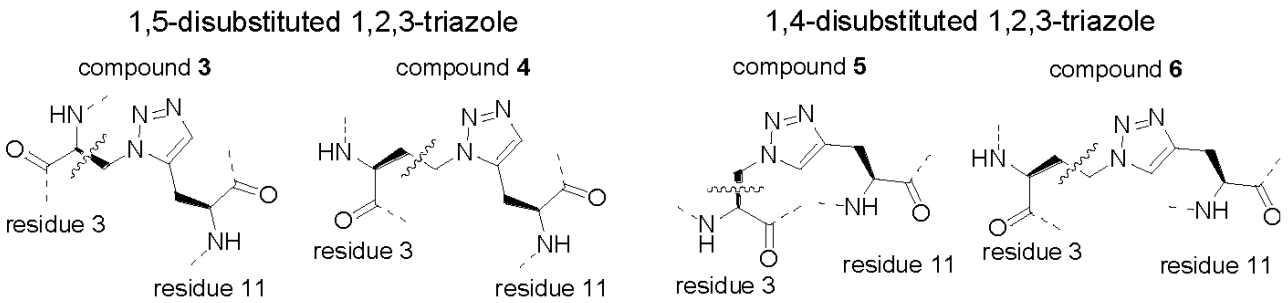
All *in silico* experiments were performed on an Intel® Core™ i7-2600 workstation using 8 virtual cores.

Graphical Content

Graphical content for Figures 1, 3, 6, S1, and S2 was generated with YASARA (www.yasara.org) and POVRay (www.povray.org).

Force Field Parameters for Triazoles within Peptidomimetics 3–6

Each triazolyl moiety was modeled manually into residue 11 of peptidomimetics **3**–**6** and then connected *via* a single C-C bond to residue 3 as depicted in Scheme S1.



Scheme S1. Affiliation of atoms within triazolyl moieties to respective residues within compounds **3**–**6**.

Supporting Figures

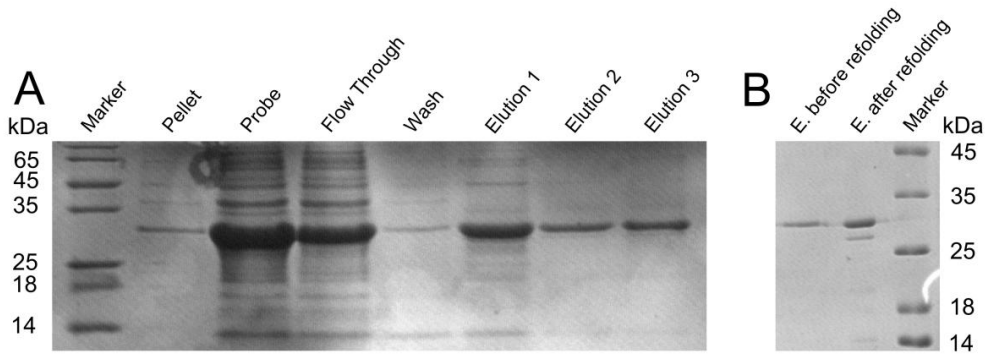


Fig S1 (A) SDS-PAGE of fractions collected using an immobilized metal ion affinity chromatography (IMAC) column upon purification of human matriptase I under denaturing conditions. Inclusion bodies were produced in *E. coli* BL21-DE3-CodonPlus-RP with the expression vector pET42dest-His-hMatI(cd)596-855 and dissolved in buffer 1 (50 mM Tris-HCl, 100 mM NaCl, 1 mM 2-mercaptoethanol, 6 M urea) after cell disruption. Elution was achieved using buffer 2 (50 mM Tris-HCl, 100 mM NaCl, 1 mM 2-mercaptoethanol, 4.5 M urea). (B) SDS-PAGE of human matriptase I before and after refolding. Refolding was achieved by 3 steps of dialysis for 4-6 hours: 1× against refolding buffer 1 (50 mM Tris-HCl, 100 mM NaCl, 1 mM 2-mercaptoethanol, 3 M urea) and then 2× against refolding buffer 2 (50 mM Tris-HCl, 100 mM NaCl, 1 mM 2-mercaptoethanol).

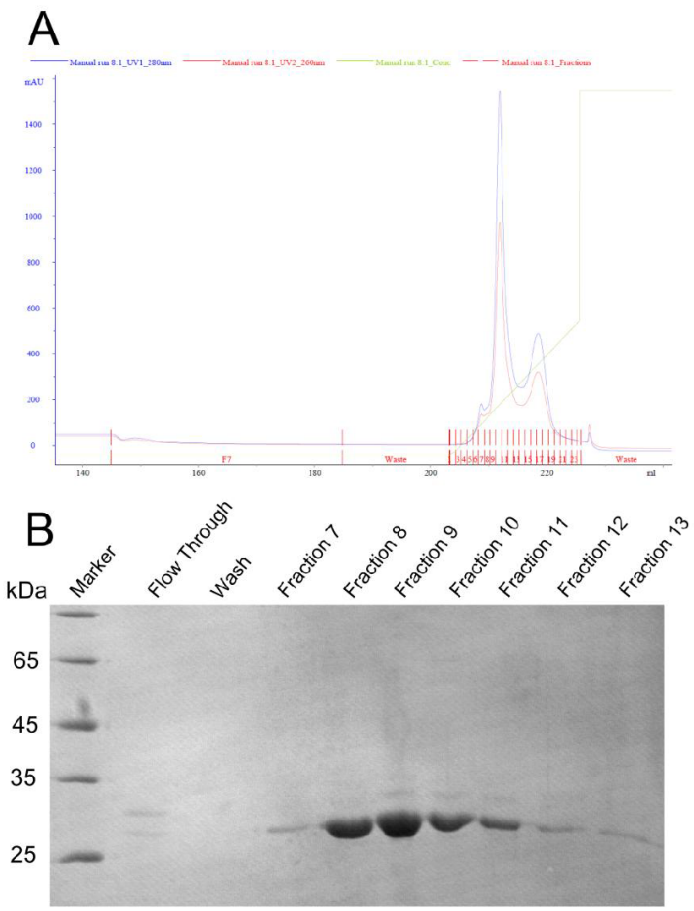


Fig S2 (A) FPLC trace of refolded/autocatalytically activated human matriptase I using an anion-exchange chromatography column (HiTrap Q HP, GE Healthcare) with detection at 260 nm (red) and 280 nm (blue). Sodium chloride was removed from protein solution before FPLC *via* dialysis against 50 mM Tris-HCl pH 8. Target protease was eluted by an increasing sodium chloride gradient 0→500 mM in 50 mM Tris-HCl (green) at a flow rate of 1 mL/min. Collected fractions are indicated (red lines). (B) SDS-PAGE of the collected fractions.

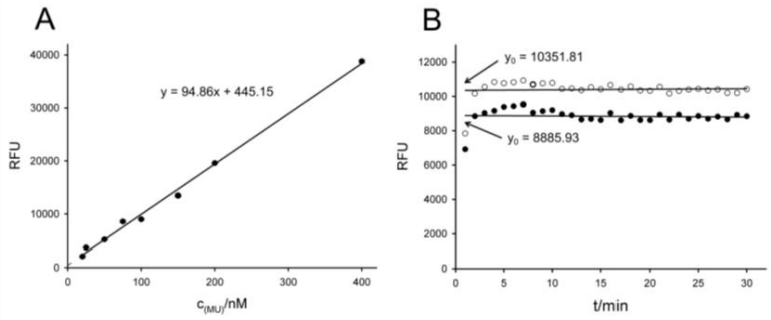


Fig S3 Analysis of proteolytic activity of the purified matriptase. (A) Calibration curve (linear regression) for fluorophor 4-methylumbelliferone (MU; Sigma) recorded at 465 nM using eight different concentrations and an excitation wavelength of 360 nM. (B) Active-site titration of purified matriptase using fluorogenic 4-methylumbelliferyl-*p*-guanidinobenzoate (MUGB; 1µM; Sigma) in buffer (50 mM Tris/HCl, 150 mM NaCl, 0.01% (v/v) Triton X-100, 0.01% (w/v) sodium azide, pH 7.6). Mean values of two independent measurements over 30 minutes (white circle 250 nM and black circle 200 nM of purified protease) y-intercept of linear regression is given. Protease activity was determined as 43 % of total protein concentration.

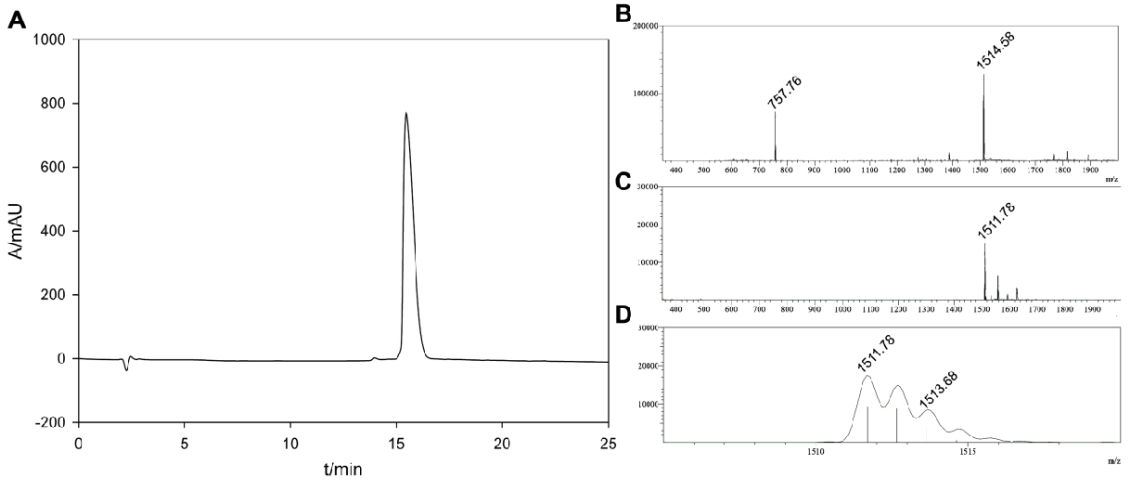


Fig S4 (A) RP-HPLC trace of bicyclic SFTI-1 (**1**). Conditions: 18 % acetonitrile over 2 min followed by 18→40.5 % acetonitrile in 0.1 % TFA over 20 min at flow rate 1 mL/min. (B) ESI mass spectrum of **1** (positive polarization). (C) ESI mass spectrum of **1** (negative polarization). (D) Section (m/z 1505-1520) of negative ESI mass spectrum showing isotopic pattern of $[M-H]^-$ signal of **1**.

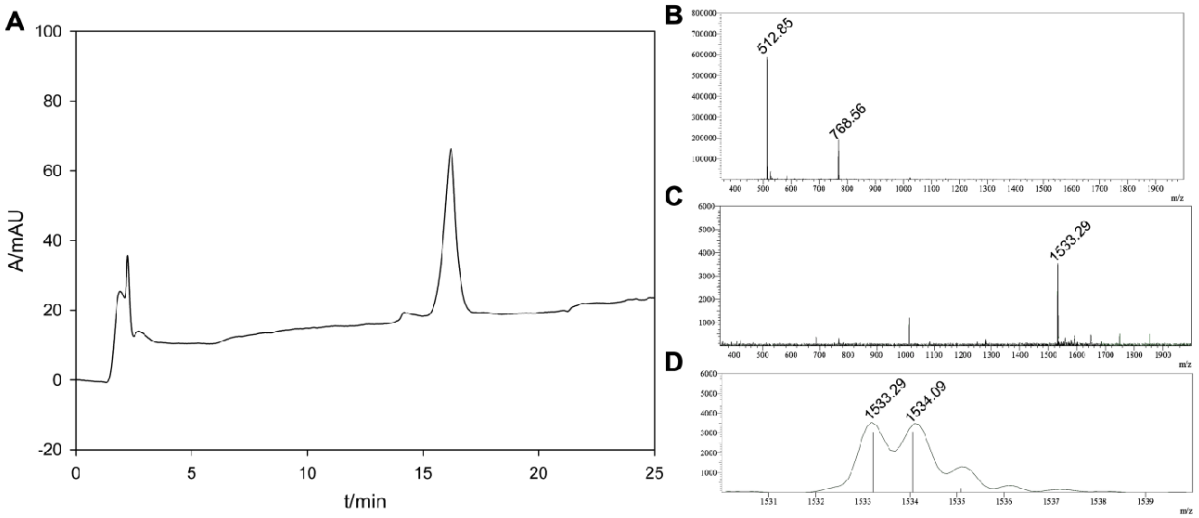


Fig S5 (A) RP-HPLC trace of peptidomimetic inhibitor (**3**). Conditions: 18 % acetonitrile over 2 min followed by 18→40.5 % acetonitrile in 0.1 % TFA over 20 min at flow rate 1 mL/min. (B) ESI mass spectrum of **3** (positive polarization). (C) ESI mass spectrum of **3** (negative polarization). (D) Section (m/z 1530-1540) of negative ESI mass spectrum showing isotopic pattern of $[M-H]^-$ signal of **3**.

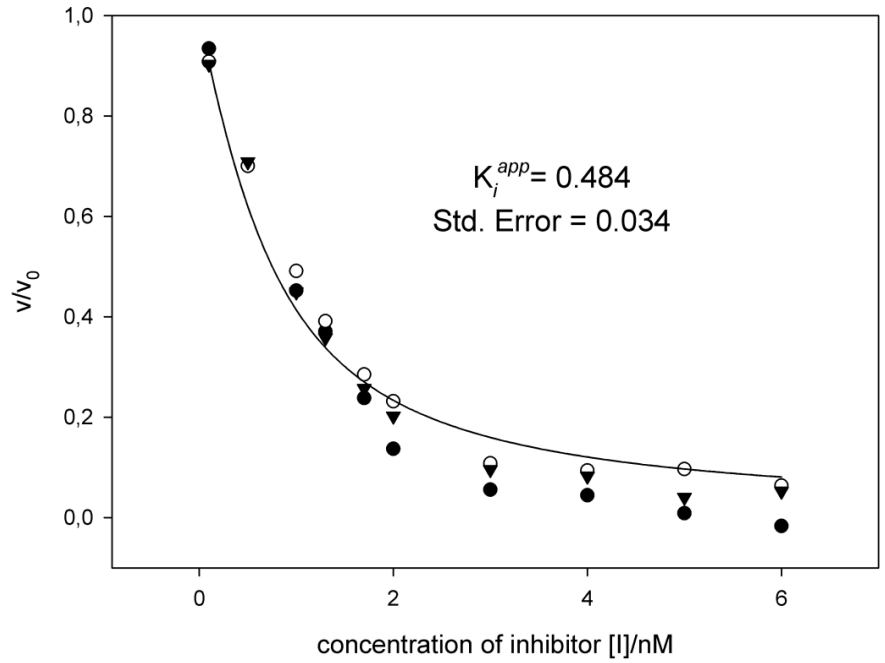


Fig. S6 Plottet kinetic data for the inhibition of the proteolytic activity of trypsin by bicyclic inhibitor **1** and resulting curve for the global non-linear fit of equation (3) onto the three sets of experimental data (1: white circle, 2: black circle, 3: black triangle). Determined apparent inhibition constant K_i^{app} and the standard error of the fit are given in nM.

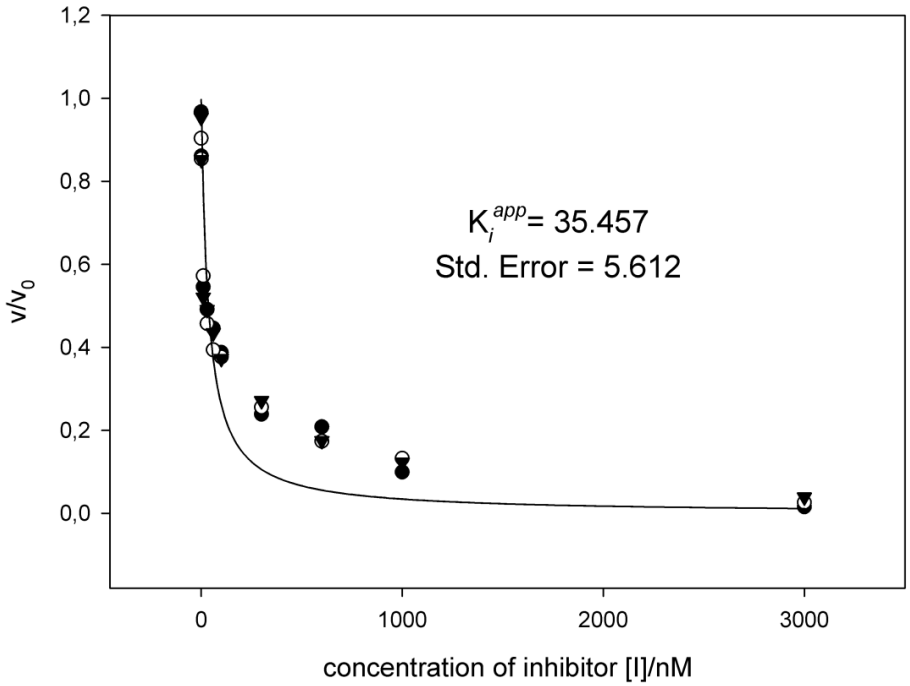


Fig. S7 Plottet kinetic data for the inhibition of the proteolytic activity of trypsin by peptidomimetic inhibitor **3** and resulting curve for the global non-linear fit of equation (3) onto the three sets of experimental data (1: white circle, 2: black circle, 3: black triangle). Determined apparent inhibition constant K_i^{app} and the standard error of the fit are given in nM.

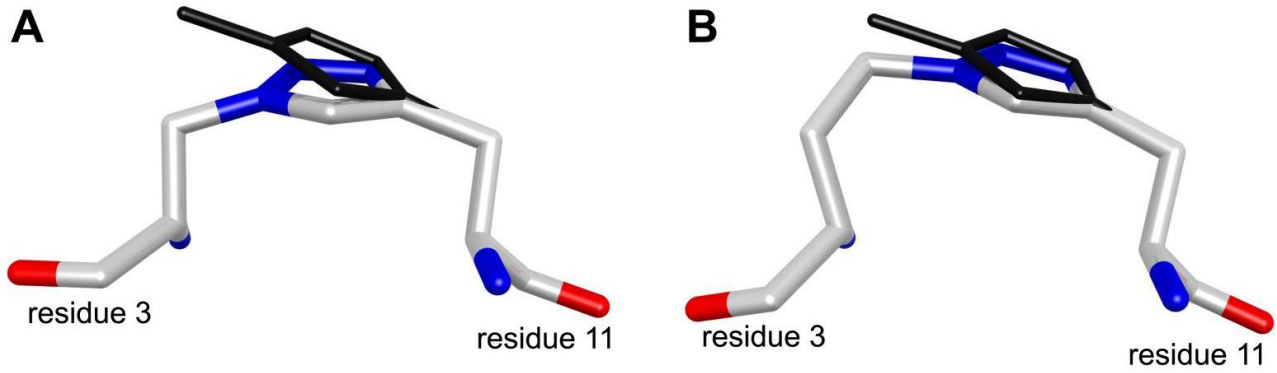


Fig S8 Bent triazolyl structures within the macrocyclization motif of compounds **5** (A) and **6** (B) resulting after modeling and a singular energy minimization step in an overlay with a corresponding planar triazole structure (thin, black). Blue: nitrogen, red: oxygen, white: carbon, hydrogen omitted for clarity, only residues 3 and 11 are shown.

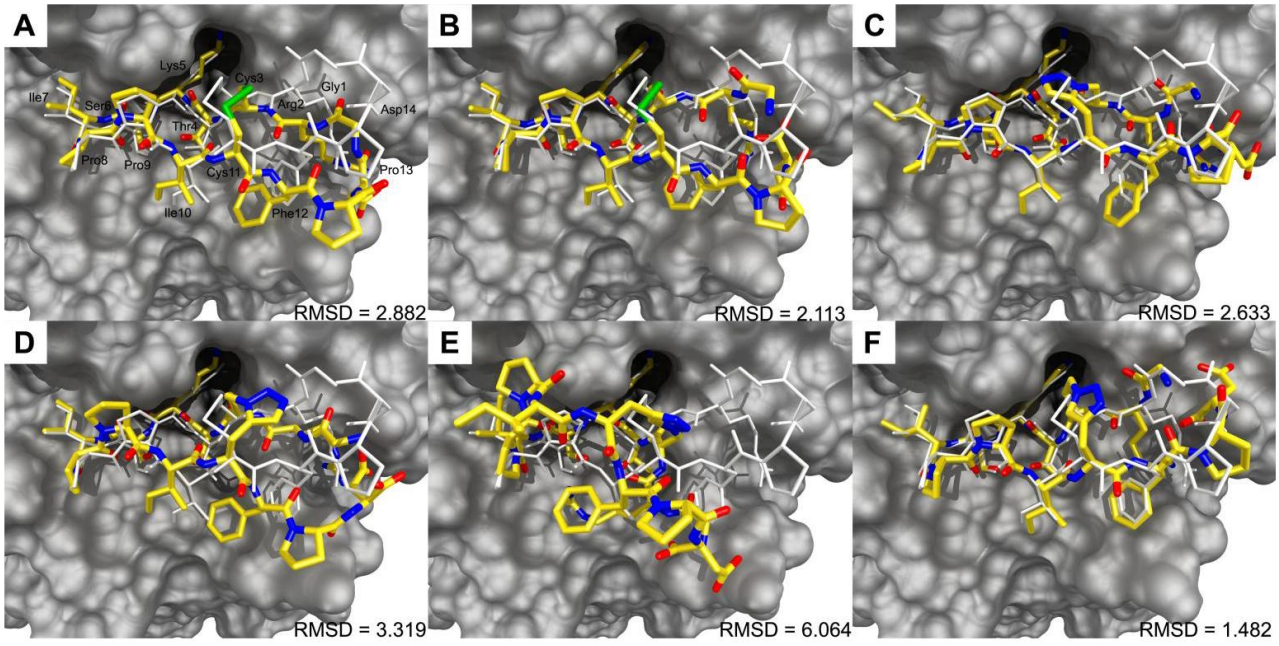


Fig. S9 Predicted structures of compounds **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), and **6** (F) in complex with trypsin (grey surface) as an overlay with reported crystal structure 1SFI (reference 2 from the main text) (white sticks). Blue: nitrogen, green: sulfur, red: oxygen, yellow: carbon, hydrogen is omitted for clarity. Measured RMSD values for inhibitor backbones compared to 1SFI are given in Å.

Supporting Tables

Table S1 Determined K_i for compounds **1-6** against trypsin at pH 7.6 using equations (1) and (4 or 5)[a,b] (K_i^1), (2) and (6) (K_i^2), as well as (3) and (6) (K_i^3).

Entry	K_i^1/nM	K_i^2/nM	K_i^3/nM
1	0.06[a]	0.07	0.07
2	0.21[a]	0.21	0.21
3	5.08[b]	5.07	5.07
4	0.34[a]	0.34	0.34
5	272.74[b]	272.72	272.73
6	106.12[b]	106.09	106.09

[a]Equation (5) was used. [b]Equation (4) was used.

Table S2 Determined K_i for compounds **1-6** against matriptase at pH 7.6 using equations (1) and (4) (K_i^1), (2) and (6) (K_i^2), as well as (3) and (6) (K_i^3).

Entry	K_i^1/nM	K_i^2/nM	K_i^3/nM
1	1100.1	1099.9	1099.9
2	702.8	702.5	702.5
3	1236.9	1236.7	1236.7
4	12930.0	12929.1	12929.8
5	285258.8	284749.5	284748.4
6	94092.1	94084.5	94096.2

Table S3 Determined K_i for compounds **1** & **2** against matriptase at pH 8.5 using equations (1) and (4) (K_i^1), (2) and (6) (K_i^2), as well as (3) and (6) (K_i^3).

Entry	K_i^1/nM	K_i^2/nM	K_i^3/nM
1	147.5	147.4	147.4
2	100.3	100.2	100.2

Table S4 Assignment of atom types for investigated 1,5-disubstituted 1,2,3-triazoles and 1,4-disubstituted 1,2,3-triazoles.

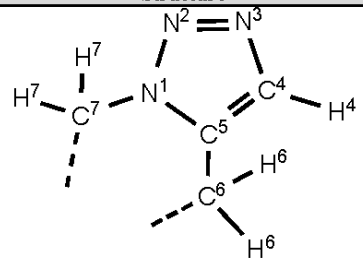
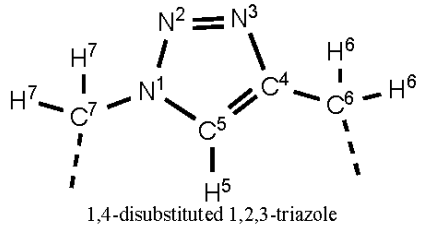
Structure	Atom	Atom type
 1,5-disubstituted 1,2,3-triazole	N^1	N/
	N^2	N(
	N^3	N)
	C^4	C/
	C^5	C(
	C^6	C2
	C^7	C2
	H^4	H4
	H^6	HC
	H^7	HC
 1,4-disubstituted 1,2,3-triazole	N^1	N$
	N^2	N%
	N^3	N&
	C^4	C$
	C^5	C%
	C^6	C2
	C^7	C2
	H^5	H4
	H^6	HC
	H^7	HC

Table S5 Bond parameters.

Atom1-Atom2	Force constant/(kcal/(mol×Å^2))	Equilibrium distance/Å
C$-C%	596.25000	1.377
C$-N&	596.25000	1.363
C2-C$	396.25000	1.509
C%-H4	458.75000	0.949
C%-N$	596.25000	1.366
N$-C2	632.25000	1.467
N&-N%	596.25000	1.293
N%-N$	596.25000	1.363
C(-C/	596.25000	1.377
C/-N)	596.25000	1.354
C2-C(	396.25000	1.479
C/-H4	458.75000	0.949
C(-N/	596.25000	1.365
N/-C2	632.25000	1.470
N)-N(	596.25000	1.322
N(-N/	596.25000	1.358
C2-h1	425.00000	1.093

Table S6 Angle parameters.

Atom1-Atom2-Atom3	Force constant/(kcal/(mol×rad^2))	Equilibrium angle/degrees
C2-C$-C%	373.02300	125.684
C2-C$-N&	373.02300	126.776
C$-C%-H4	95.500000	126.385
C$-C%-N$	373.02300	107.231
C$-N&-N%	373.02300	107.954
C%-N$-C2	373.02300	131.073
C%-C$-N&	373.02300	107.540
N&-N%-N$	373.02300	111.267
C%-N$-N%	373.02300	105.954

H4-C%-N$	95.500000	126.385
N%-N$-C2	373.02300	122.982
C$-C2-C1	64.700000	108.100
C$-C2-HC	47.200000	110.860
C1-C2-N$	65.800000	112.590
N$-C2-h1	49.900000	109.450
C2-C2-N$	65.800000	112.590
HC-C2-N$	49.900000	109.500
C2-C(-C/	373.02300	129.650
C2-C(-N/	373.02300	126.517
C(-C/-H4	95.500000	125.288
C(-C/-N)	373.02300	109.412
C/-N)-N(	373.02300	108.718
C(-N/-C2	373.02300	129.074
N)-N(-N/	373.02300	107.118
C(-N/-N(	373.02300	111.158
H4-C/-N)	95.500000	125.300
N(-N/-C2	373.02300	119.767
C/-C(-N/	373.02300	103.833
C(-C2-C1	64.700000	108.100
C(-C2-HC	47.200000	110.860
C1-C2-N/	65.800000	112.590
N/-C2-h1	49.900000	109.450
C2-C2-N/	65.800000	112.590
HC-C2-N/	49.900000	109.500
h1-C2-h1	39.200000	109.550
HC-C1-N	49.800000	109.500
C -C1-HC	47.200000	109.680
C1-C2-h1	46.400000	110.070
h1-C2-h1	39.200000	109.550
HC-C1-N	49.800000	109.500
C -C1-HC	47.200000	109.680

Table S7 Dihedral angle parameters.

Atom1-Atom2-Atom3-Atom4	Bond paths	Force constant/(kcal/mol)	Phase angle/degrees	periodicity
C2-C$-N&-N%	2	11.50000	180.00	2
C%-C$-N&-N%	2	11.50000	180.00	2
C2-C$-C%-N$	4	23.69000	180.00	2
C2-C$-C%-H4	4	23.69000	180.00	2
N&-C$-C%-N$	4	23.69000	180.00	2
N&-C$-C%-H4	4	23.69000	180.00	2
C$-C%-N$-C2	2	11.50000	180.00	2
C$-C%-N$-N%	2	11.50000	180.00	2
H4-C%-N$-C2	2	11.50000	180.00	2
H4-C%-N$-N%	2	11.50000	180.00	2
C2-N$-N%-N&	2	9.600	180.00	2
C%-N$-N%-N&	2	9.600	180.00	2
C$-N&-N%-N$	1	4.000	180.00	2
X -C$-C2-X	6	0.000	0.000	3
X -C2-N$-X	6	0.000	0.000	3
H4-C/-N)-N(	2	11.50000	180.00	2
C(-C/-N)-N(	2	11.50000	180.00	2
N/-C(-C/-H4	4	23.69000	180.00	2
C2-C(-C/-H4	4	23.69000	180.00	2
N/-C(-C/-N)	4	23.69000	180.00	2
C2-C(-C/-N)	4	23.69000	180.00	2
C/-C(-N/-C2	2	11.50000	180.00	2
C/-C(-N/-N(	2	11.50000	180.00	2
C2-C(-N/-C2	2	11.50000	180.00	2
C2-C(-N/-N(	2	11.50000	180.00	2
C2-N/-N(-N)	2	9.600	180.00	2
C(-N/-N(-N)	2	9.600	180.00	2
C/-N)-N(-N/	1	4.000	180.00	2
X -C2-C(-X	6	0.000	0.000	3
X -C2-N/-X	6	0.000	0.000	3

6.2. Supporting information for chapter 2.2

Contents: **Pages:**

ESI-MS	S2
RP-HPLC Data	S9
NMR-spectra	S14
Plotted Kinetic Data	S18
Additional Figures	S25

1. ESI-MS

Table S1. ESI-MS spectral data for compounds **1-22** and **42-44**.

Entry	SIM / g·mol^{-1}	$[M+H]^+$	$[M+2H]^{2+}$	$[M+3H]^{3+}$	$[M-2H]^{2-}$
1	1585.8	-	794.9 (793.9)	530.3 (529.6)	-
2	1599.8	-	801.2 (800.9)	534.6 (534.3)	-
3	1600.8	1601.9 (1601.8)	801.1 (801.4)	534.5 (534.6)	-
4	1614.8	1615.8 (1615.4)	808.2 (808.4)	539.3 (539.3)	806.3 (806.4)
5	1614.8	-	808.3 (808.4)	539.3 (539.3)	806.2 (806.4)
6	1628.8	1630.0 (1629.8)	815.4 (815.4)	544.0 (543.9)	813.2 (813.4)
7	1632.9	-	817.4 (817.5)	545.4 (545.3)	815.4 (815.5)
8	1646.9	1648.1 (1647.9)	824.5 (824.5)	550.0 (550.0)	-
9	1585.8	-	794.0 (793.9)	529.8 (529.6)	791.8 (791.8)
10	1599.8	1600.5 (1600.8)	800.9 (800.9)	534.2 (534.3)	-
11	1600.8	-	801.0 (801.4)	534.5 (534.6)	799.0 (799.4)
12	1614.8	1615.5 (1615.8)	808.2 (808.4)	539.1 (539.3)	-
13	1614.8	1615.7 (1615.8)	808.4 (808.4)	539.4 (539.3)	806.2 (806.4)
14	1628.8	1630.1 (1629.8)	815.5 (815.5)	544.0 (543.9)	813.3 (813.4)
15	1632.9	-	817.5 (817.5)	545.4 (545.3)	-
16	1646.9	1647.9 (1647.9)	824.5 (824.5)	550.0 (550.0)	-
17	1546.8	1547.7 (1547.8)	774.5 (774.4)	516.7 (516.8)	772.5 (772.4)
18	1574.8	-	789.0 (788.4)	526.4 (525.9)	786.2 (786.4)
19	1455.7	1457.0 (1456.7)	728.9 (728.9)	486.4 (486.2)	-
20	1483.8	1484.5 (1484.8)	742.8 (742.9)	495.7 (495.6)	740.8 (740.9)
21	1497.8	1498.7 (1498.8)	749.9 (749.9)	500.3 (500.3)	-
22	1521.8		761.9 (761.9)	508.3 (508.3)	508.2 (508.3)
42	1575.8	-	788.9 (788.9)	526.3 (526.3)	786.7 (786.9)
43	1564.8	1565.7 (1565.8)	783.3 (783.4)	522.6 (522.6)	781.3 (781.4)
44	1351.6	-	676.7 (676.8)	451.4 (451.5)	674.7 (674.8)

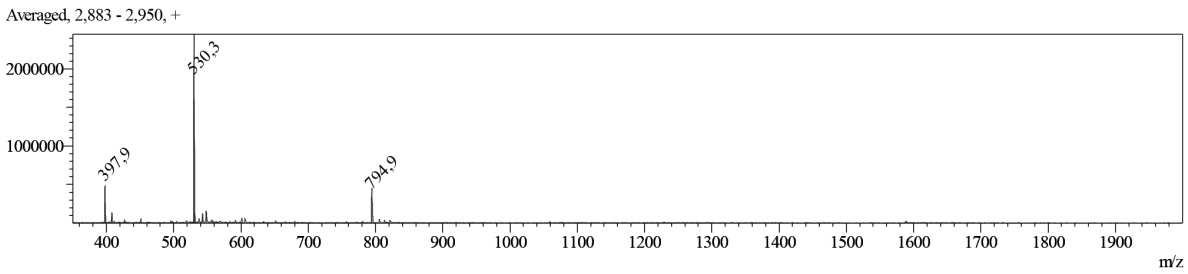


Figure S1. ESI-MS spectrum of compound 1 (positive polarization).

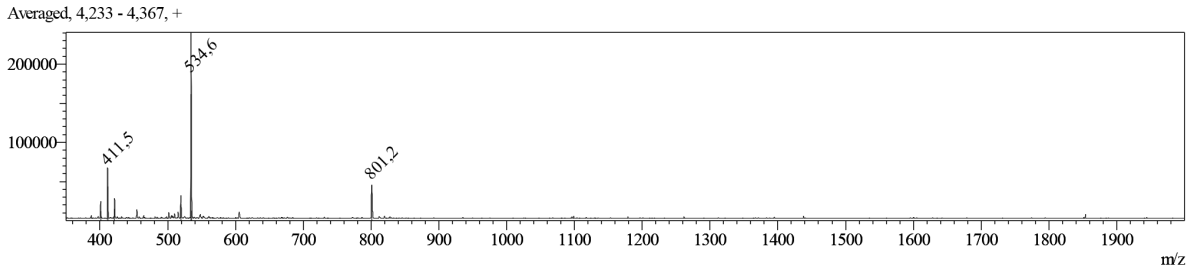


Figure S2. ESI-MS spectrum of compound **2** (positive polarization).

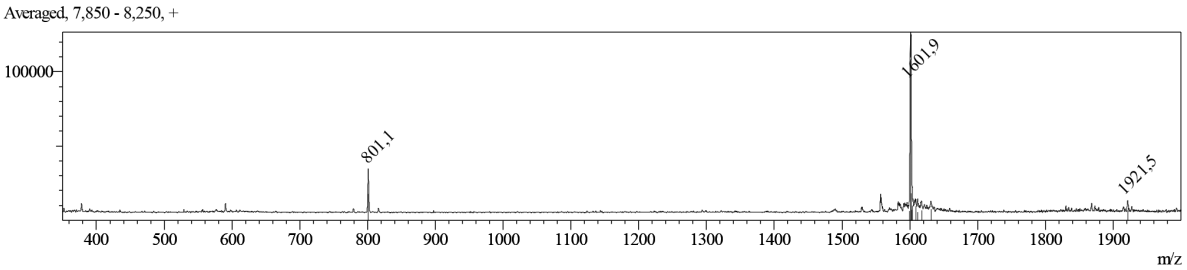


Figure S3. ESI-MS spectrum of compound **3** (positive polarization).

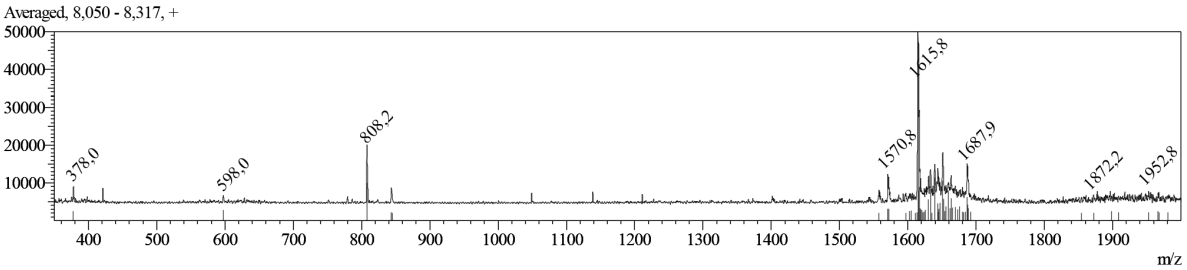


Figure S4. ESI-MS spectrum of compound **4** (positive polarization).

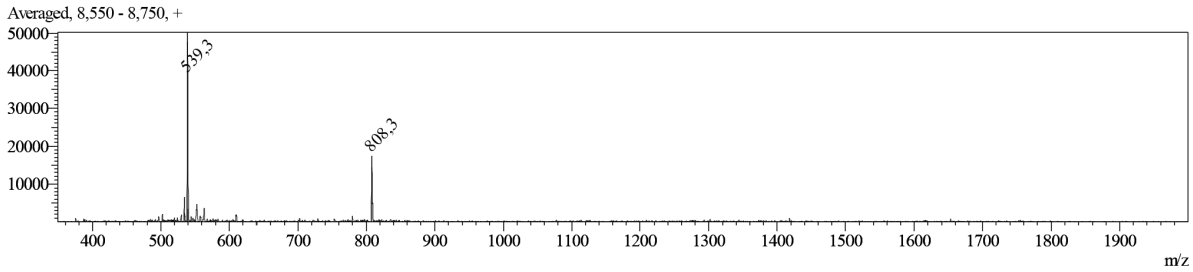


Figure S5. ESI-MS spectrum of compound **5** (positive polarization).

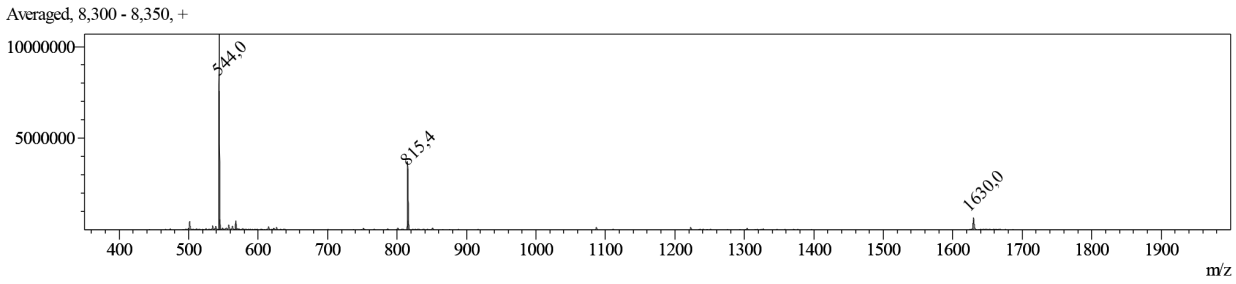


Figure S6. ESI-MS spectrum of compound **6** (positive polarization).

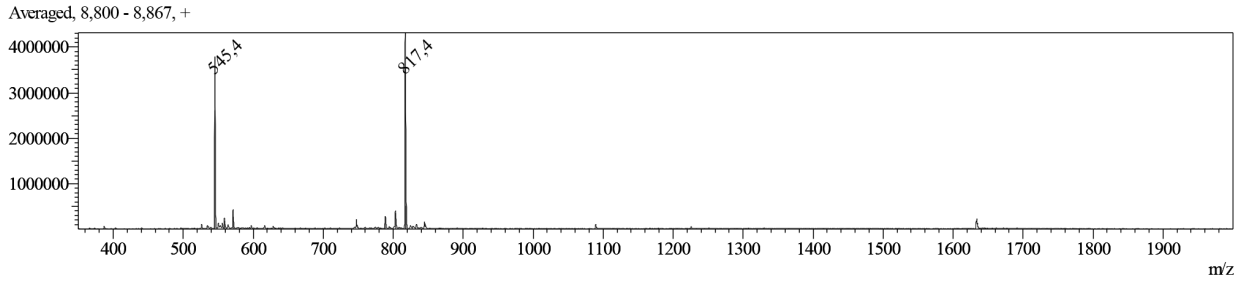


Figure S7. ESI-MS spectrum of compound **7** (positive polarization).

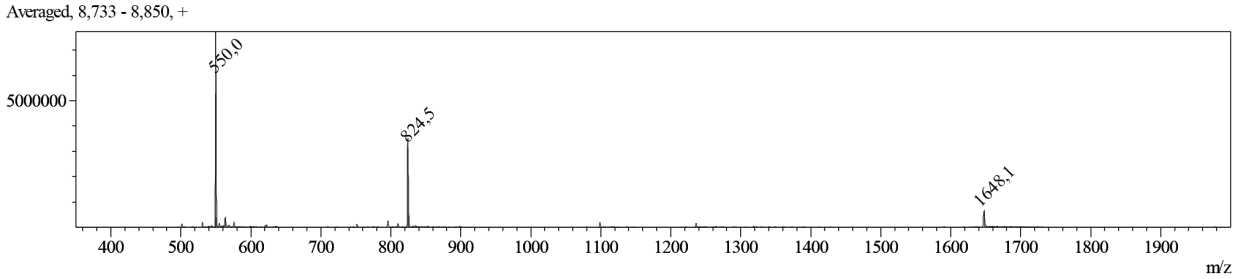


Figure S8. ESI-MS spectrum of compound **8** (positive polarization).

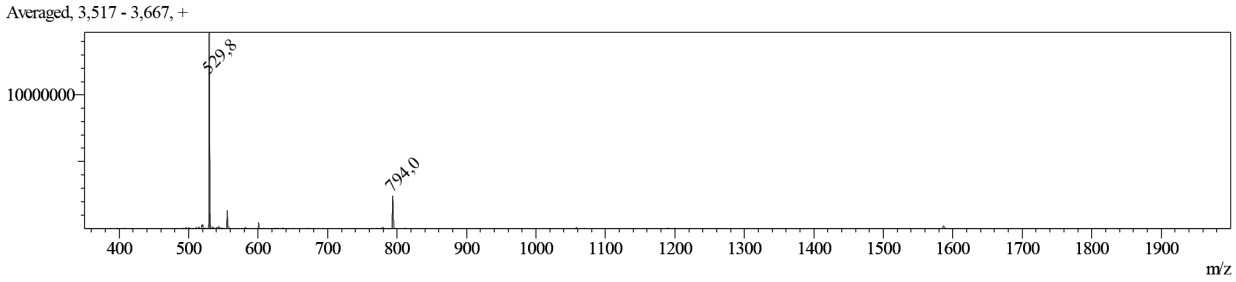


Figure S9. ESI-MS spectrum of compound **9** (positive polarization).

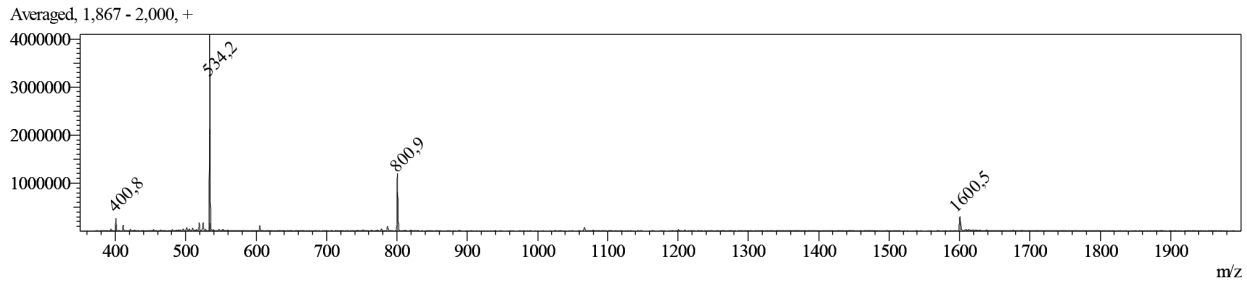


Figure S10. ESI-MS spectrum of compound **10** (positive polarization).

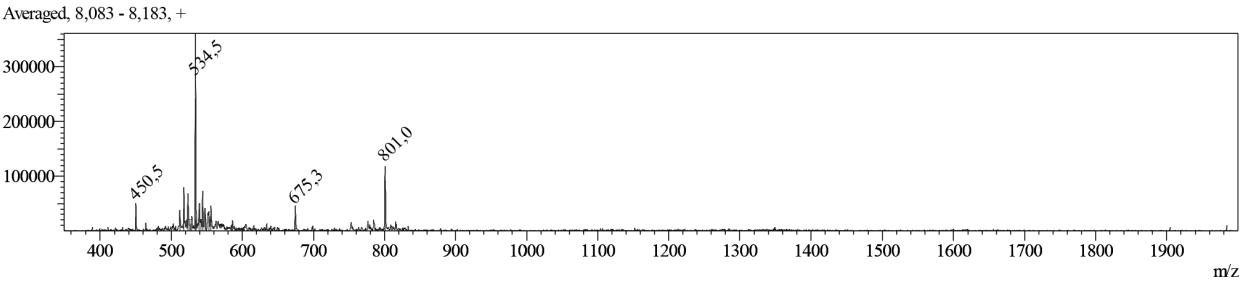


Figure S11. ESI-MS spectrum of compound **11** (positive polarization).

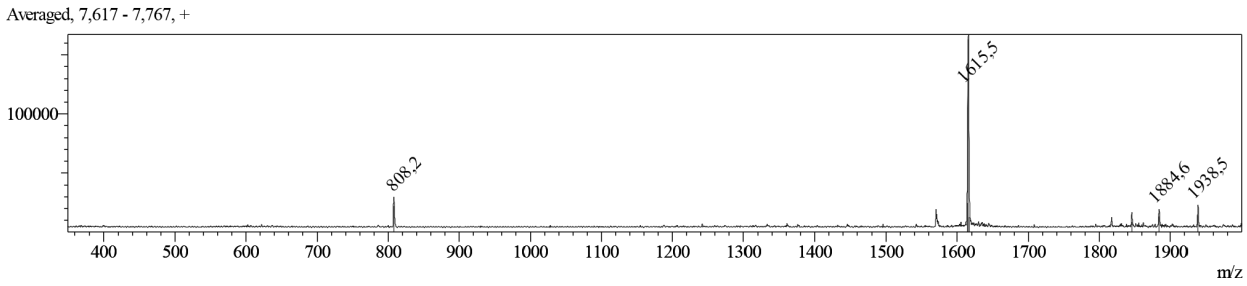


Figure S12. ESI-MS spectrum of compound **12** (positive polarization).

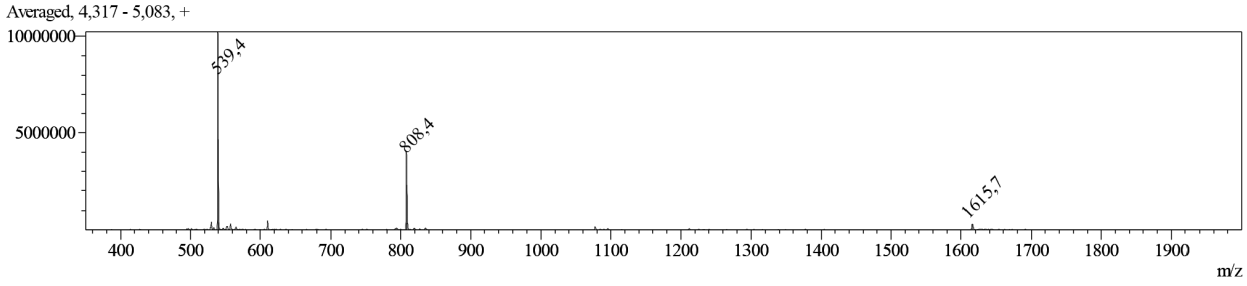


Figure S13. ESI-MS spectrum of compound **13** (positive polarization).

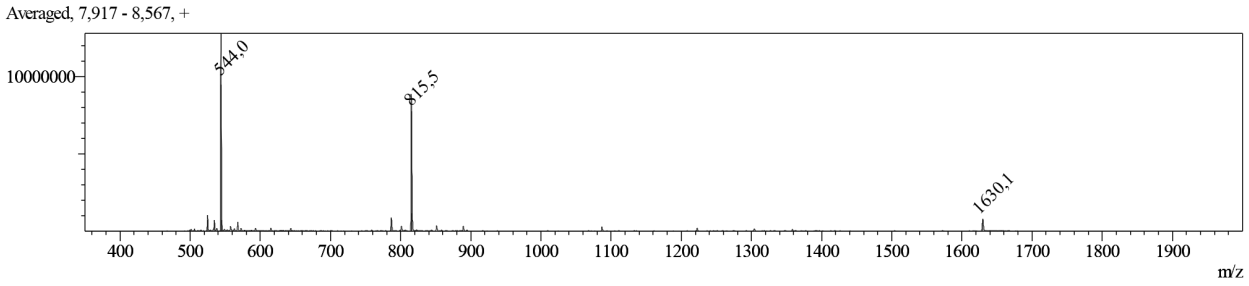


Figure S14. ESI-MS spectrum of compound **14** (positive polarization).

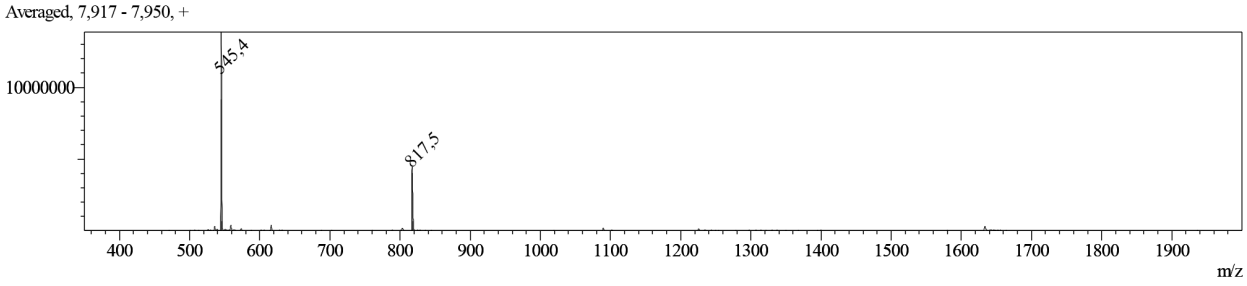


Figure S15. ESI-MS spectrum of compound **15** (positive polarization).

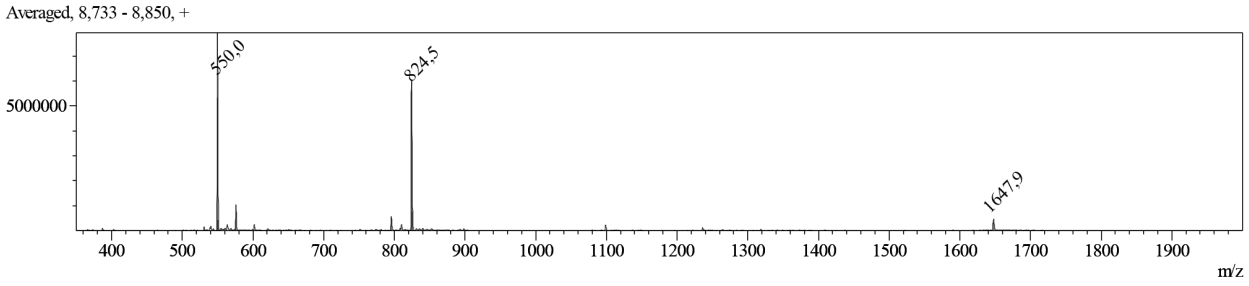


Figure S16. ESI-MS spectrum of compound **16** (positive polarization).

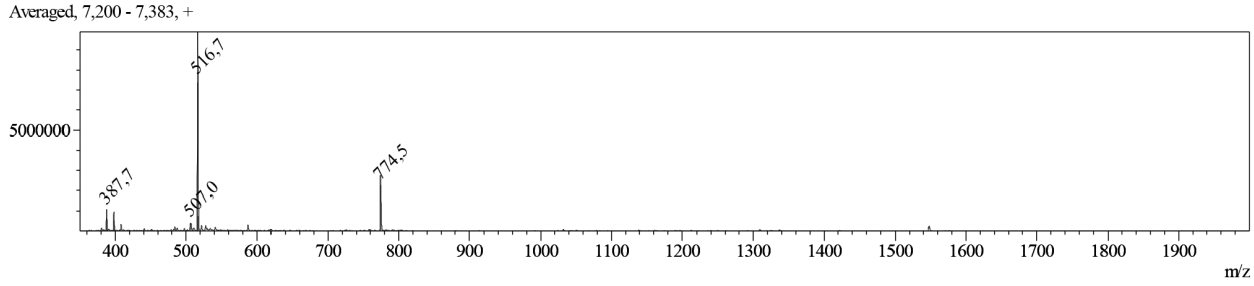


Figure S17. ESI-MS spectrum of compound **17** (positive polarization).

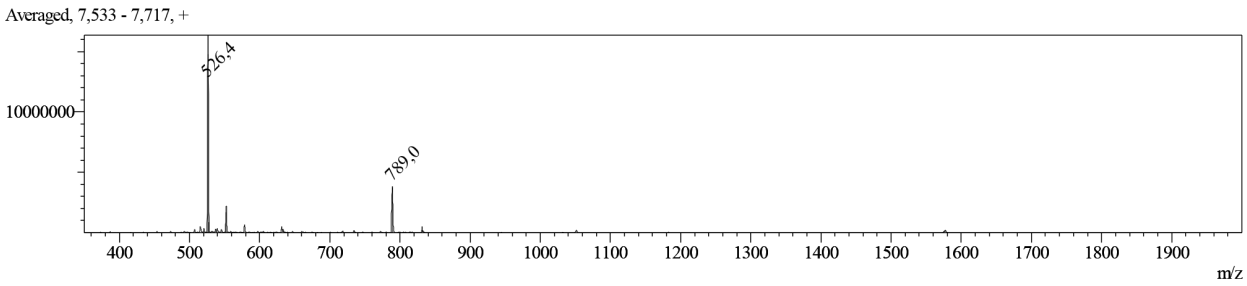


Figure S18. ESI-MS spectrum of compound **18** (positive polarization).

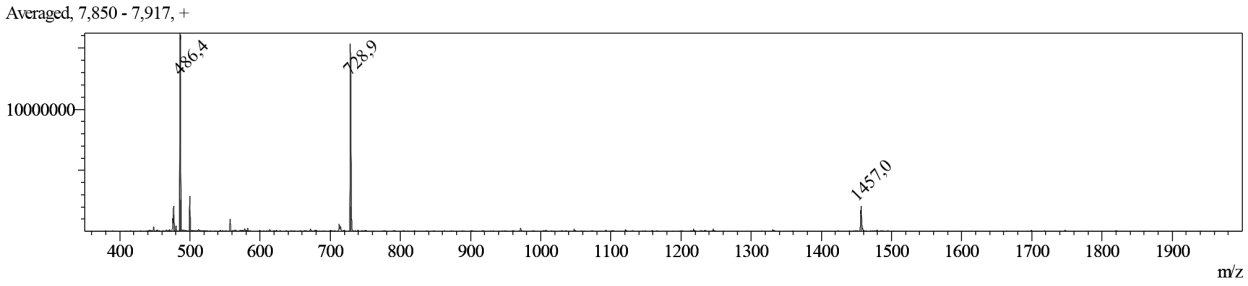


Figure S19. ESI-MS spectrum of compound **19** (positive polarization).

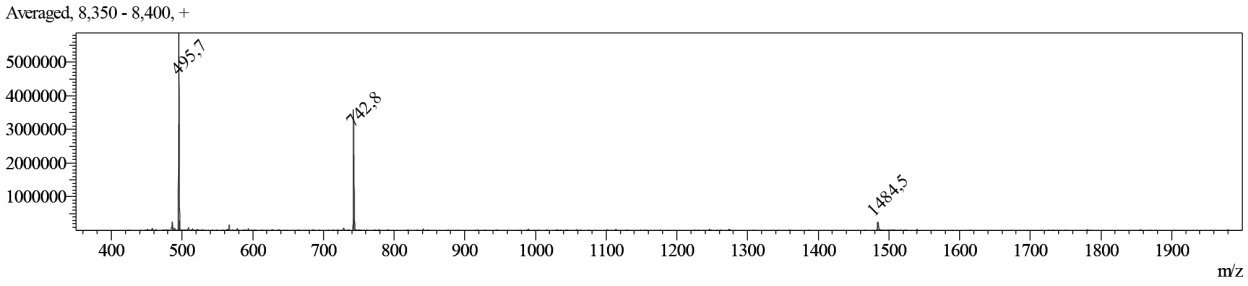


Figure S20. ESI-MS spectrum of compound **20** (positive polarization).

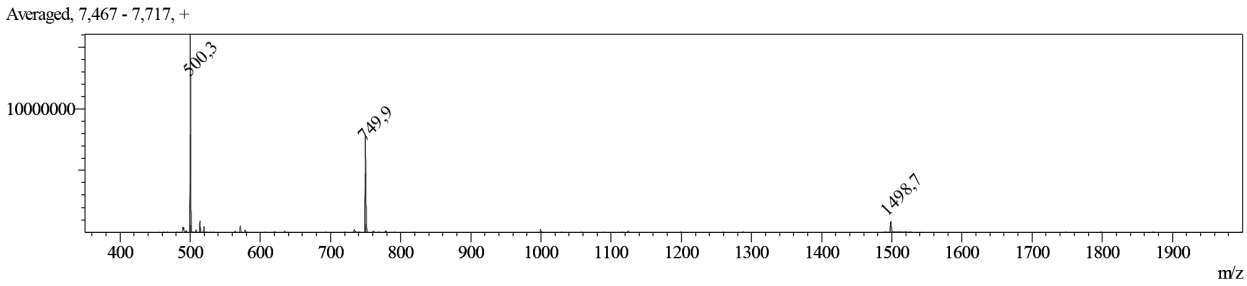


Figure S21. ESI-MS spectrum of compound **21** (positive polarization).

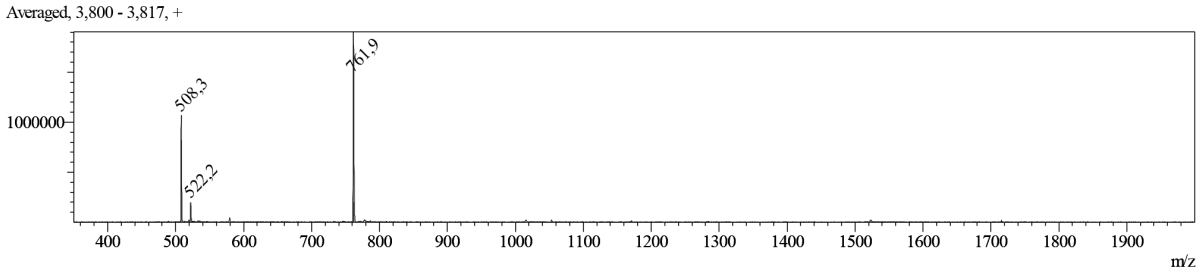


Figure S22. ESI-MS spectrum of compound **22** (positive polarization).

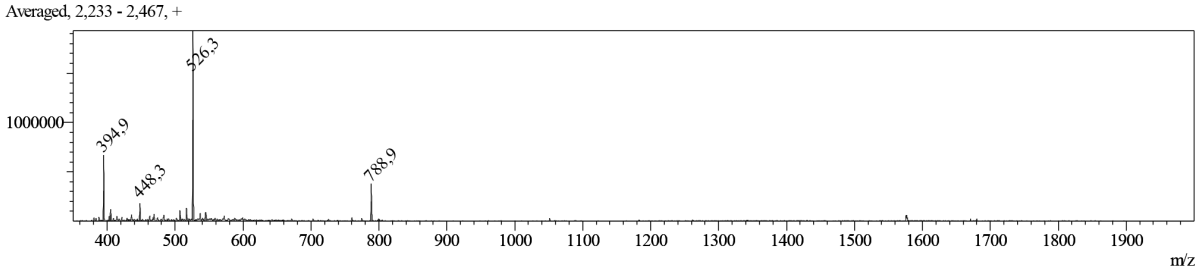


Figure S23. ESI-MS spectrum of compound **42** (positive polarization).

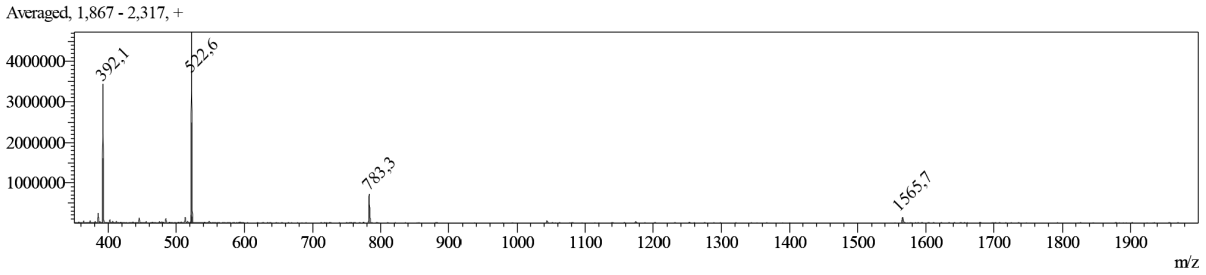


Figure S24. ESI-MS spectrum of compound **43** (positive polarization).

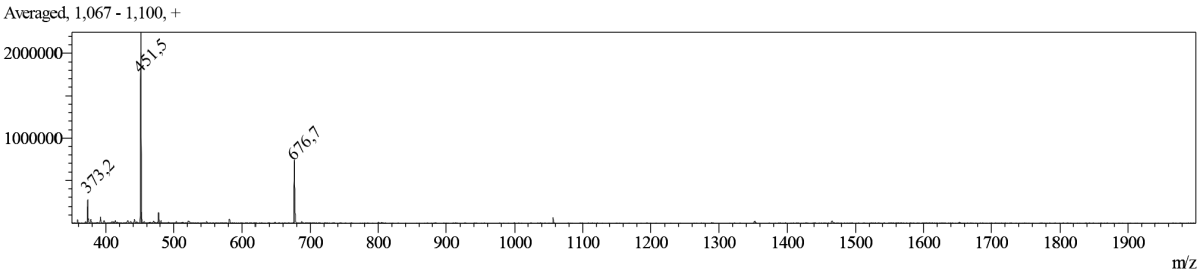


Figure S25. ESI-MS spectrum of compound **44** (positive polarization).

2. RP-HPLC Data

Table S2. Retention time of compounds **1-22** and **42-44** measured with RP-HPLC.

Entry	R_t /min	Gradient
1	12.0	18⟶40.5 %[a]
2	11.8	18⟶40.5 %[a]
3	12.5	18⟶40.5 %[a]
4	12.7	18⟶40.5 %[a]
5	13.3	18⟶40.5 %[a]
6	12.9	18⟶40.5 %[a]
7	15.9	18⟶40.5 %[a]
8	16.7	18⟶40.5 %[a]
9	7.6	18⟶40.5 %[a]
10	7.0	18⟶40.5 %[a]
11	11.6	18⟶40.5 %[a]
12	13.2	18⟶40.5 %[a]
13	11.6	18⟶40.5 %[a]
14	11.7	18⟶40.5 %[a]
15	16.5	18⟶40.5 %[a]
16	16.7	18⟶40.5 %[a]
17	18.2	0⟶36 %[b]
18	18.5	0⟶36 %[b]
19	9.8	18⟶40.5 %[a]
20	12.2	18⟶40.5 %[a]
21	13.5	18⟶40.5 %[a]
22	13.9	18⟶40.5 %[a]
42	11.5	9⟶45 %[c]
43	12.0	9⟶45 %[c]
44	11.4	9⟶45 %[c]

[a] 18 % acetonitrile over 2 min followed by 18⟶40.5 % acetonitrile in 0.1 % TFA over 20 min of flow rate 1 mL/min [b] 0 % acetonitrile over 2 min followed by 0⟶36 % acetonitrile in 0.1 % TFA over 20 min of flow rate 1 mL/min. [c] 9 % acetonitrile over 2 min followed by 9⟶45 % acetonitrile in 0.1 % TFA over 20 min of flow rate 1 mL/min.

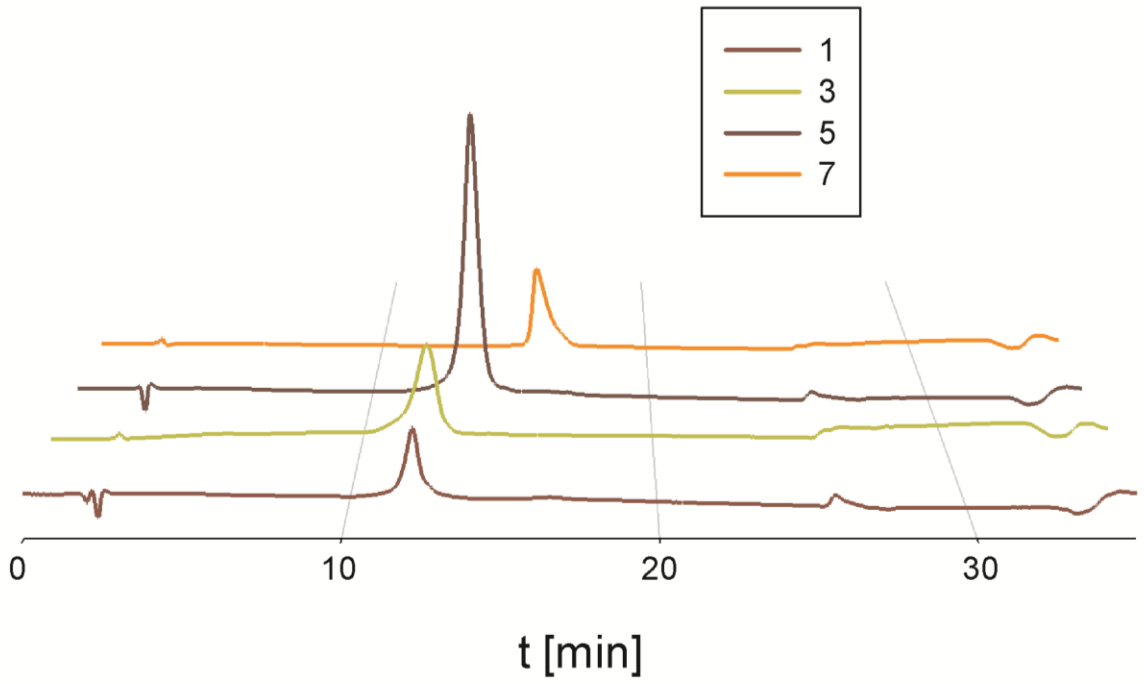


Figure S26. HPLC chromatogram of purified **1**, **3**, **5** and **7** recorded at 220 nm. Gradient: 18→40.5 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

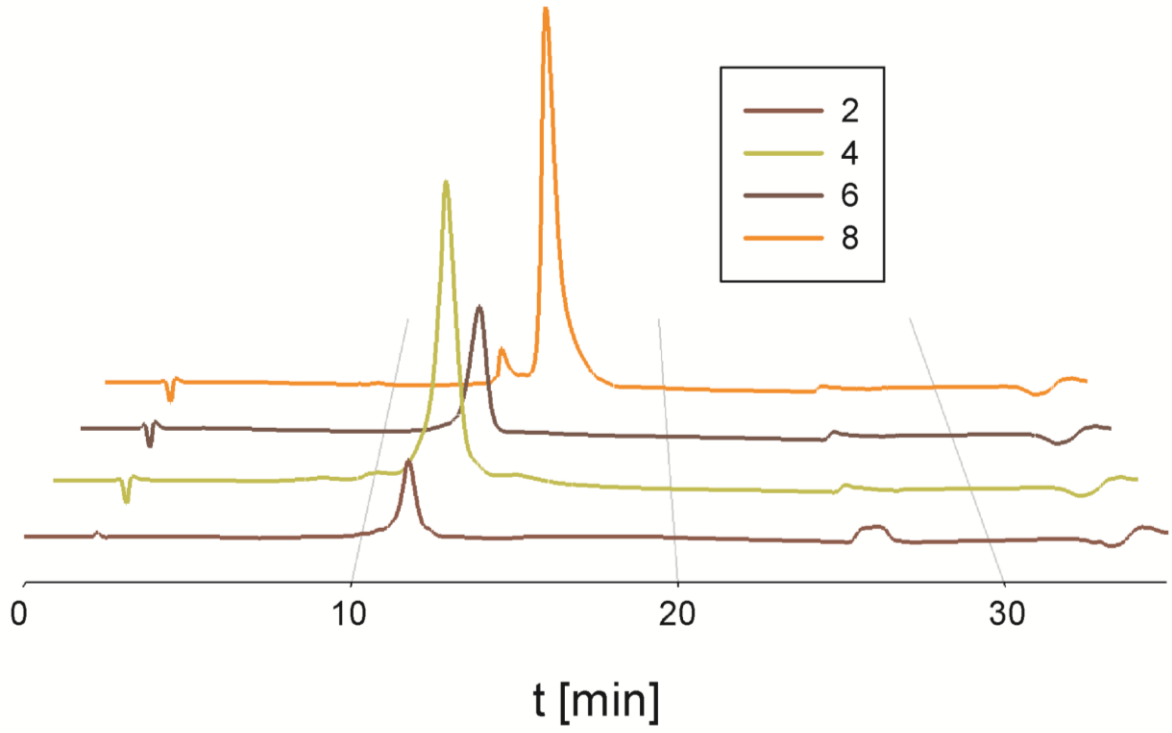


Figure S27. HPLC chromatogram of purified **2**, **4**, **6** and **8** recorded at 220 nm. Gradient: 18→40.5 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

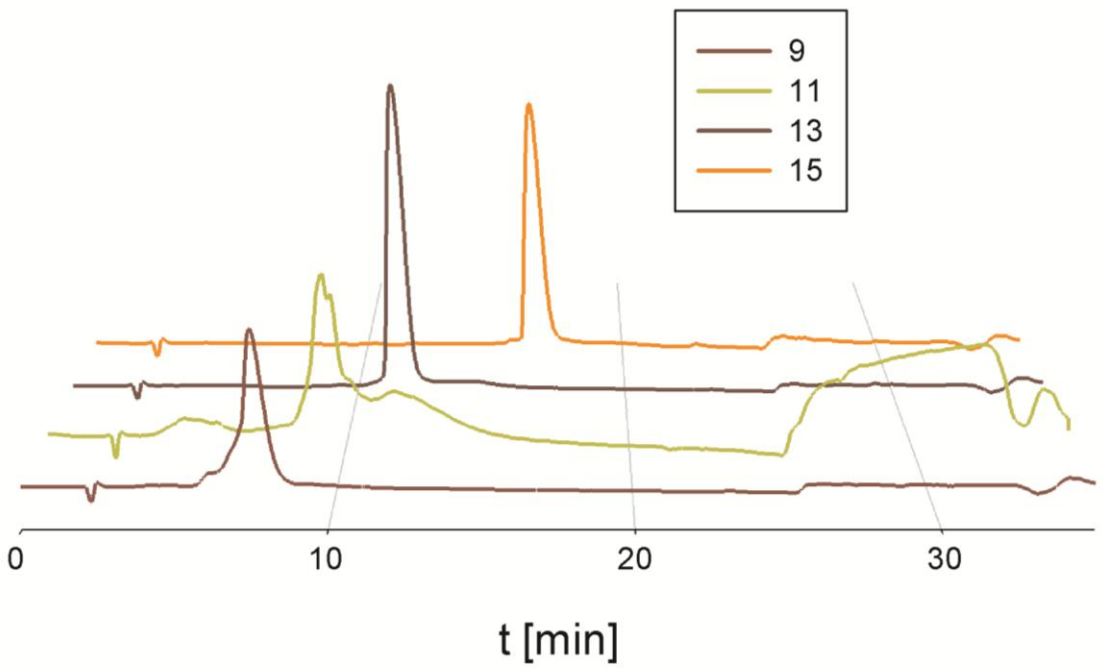


Figure S28. HPLC chromatogram of purified **9**, **11**, **13** and **15** recorded at 220 nm. Gradient: 18→40.5 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

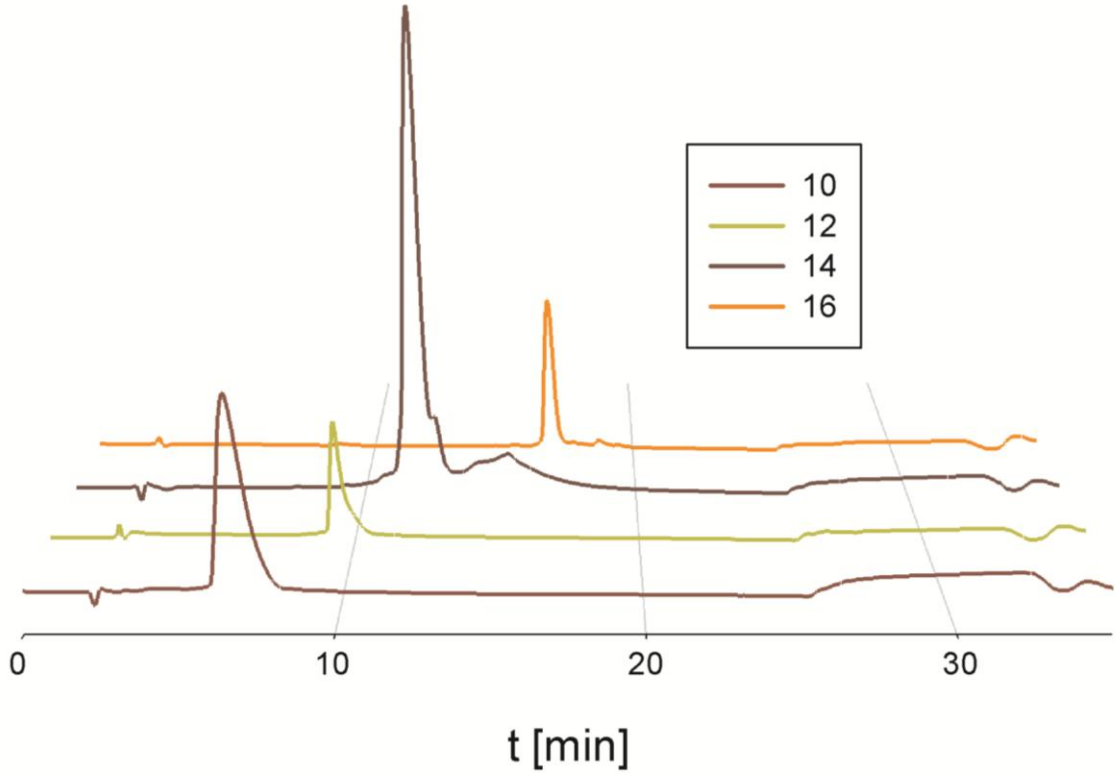


Figure S29. HPLC chromatogram of purified **10**, **12**, **14** and **16** recorded at 220 nm. Gradient: 18→40.5 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

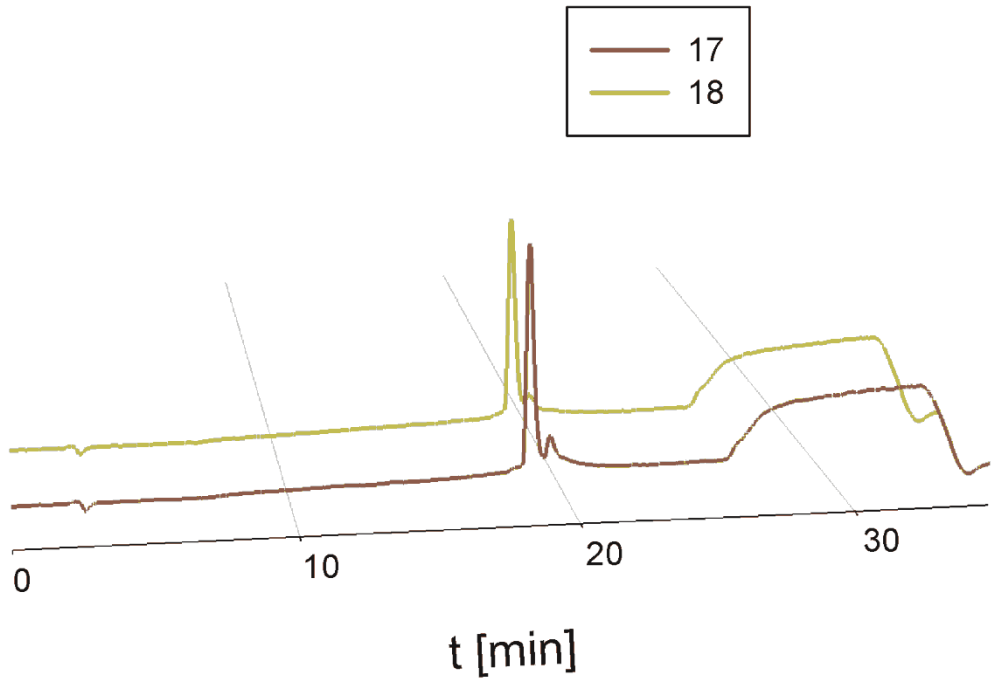


Figure S30. HPLC chromatogram of purified **17** and **18** recorded at 220 nm. Gradient: 0→36 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

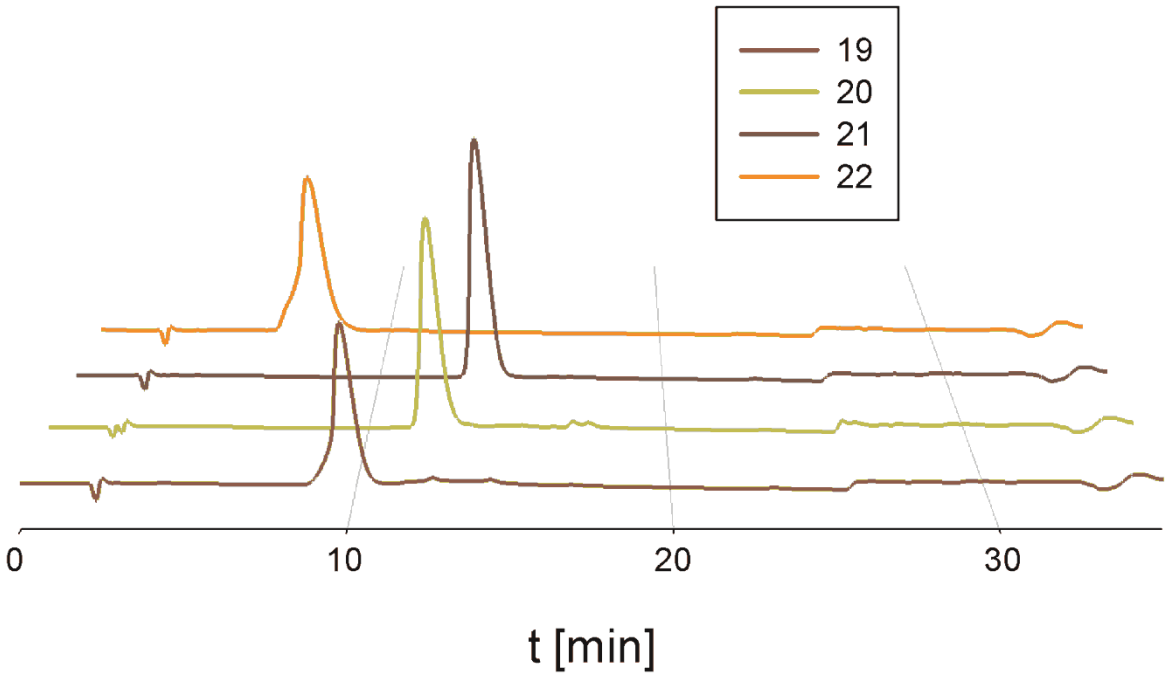


Figure S31. HPLC chromatogram of purified **19**, **20**, **21** and **22** recorded at 220 nm. Gradient: 18→40.5 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

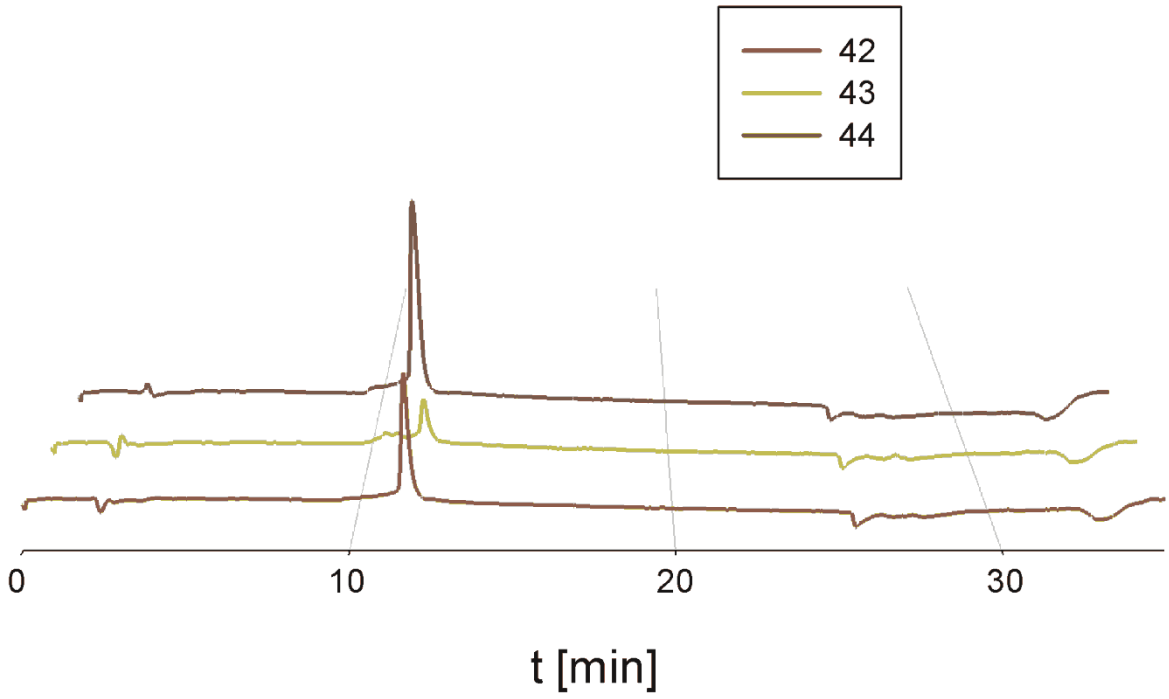


Figure S32. HPLC chromatogram of purified **42**, **43** and **44** recorded at 220 nm. Gradient: 9→45 % acetonitrile in 0.1% aq. TFA over 20 minutes at flow rate 1 mL/min.

3. NMR-spectra

Compound 43

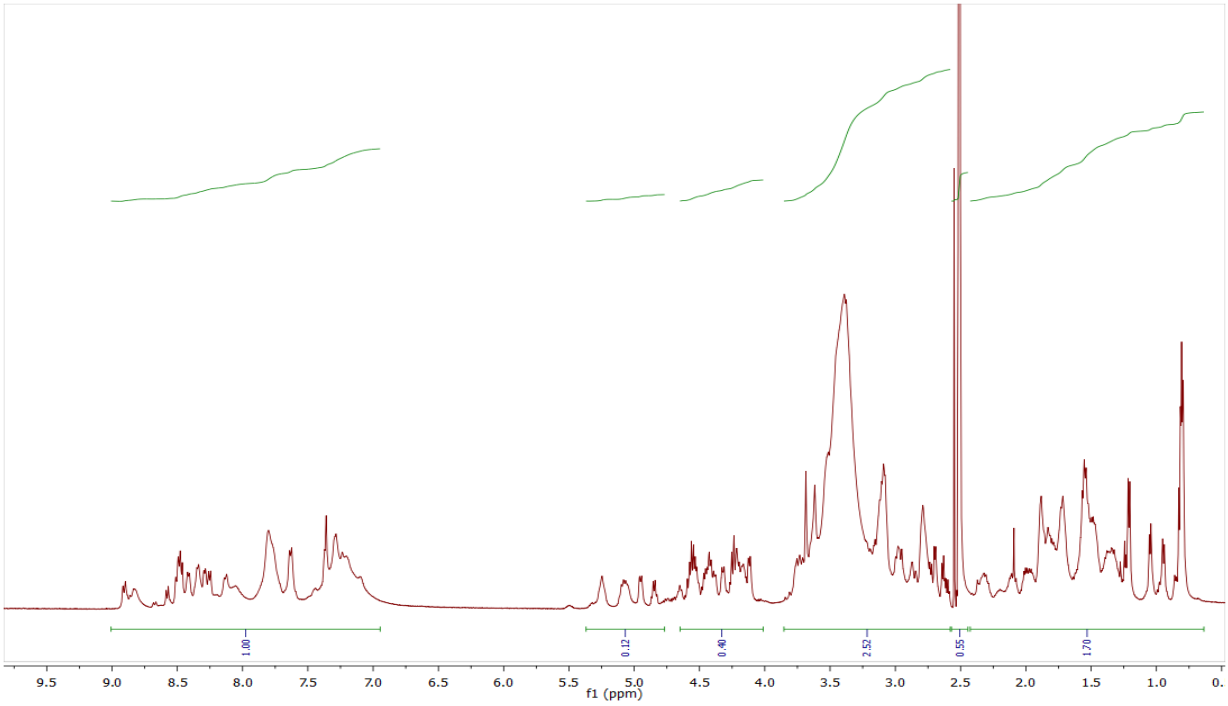


Figure S33. ^{1}H-NMR spectrum of compound **43.**

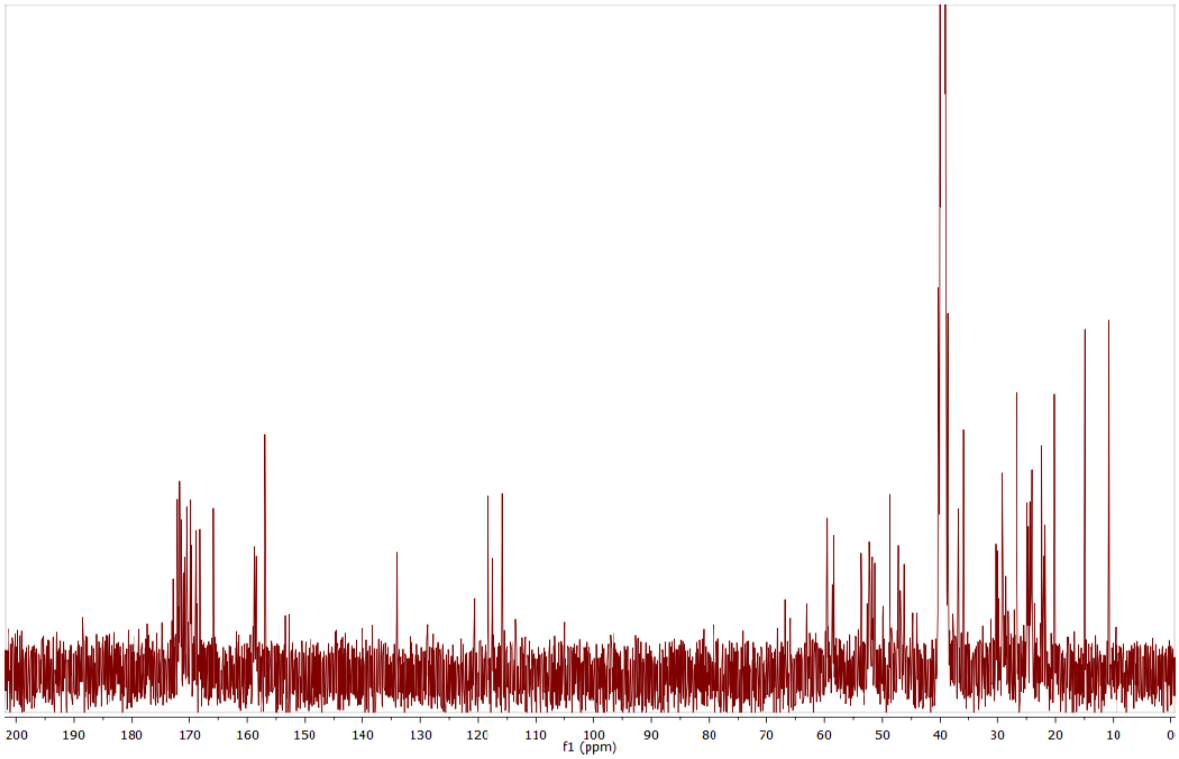


Figure S34. ^{13}C-NMR spectrum of compound **43.**

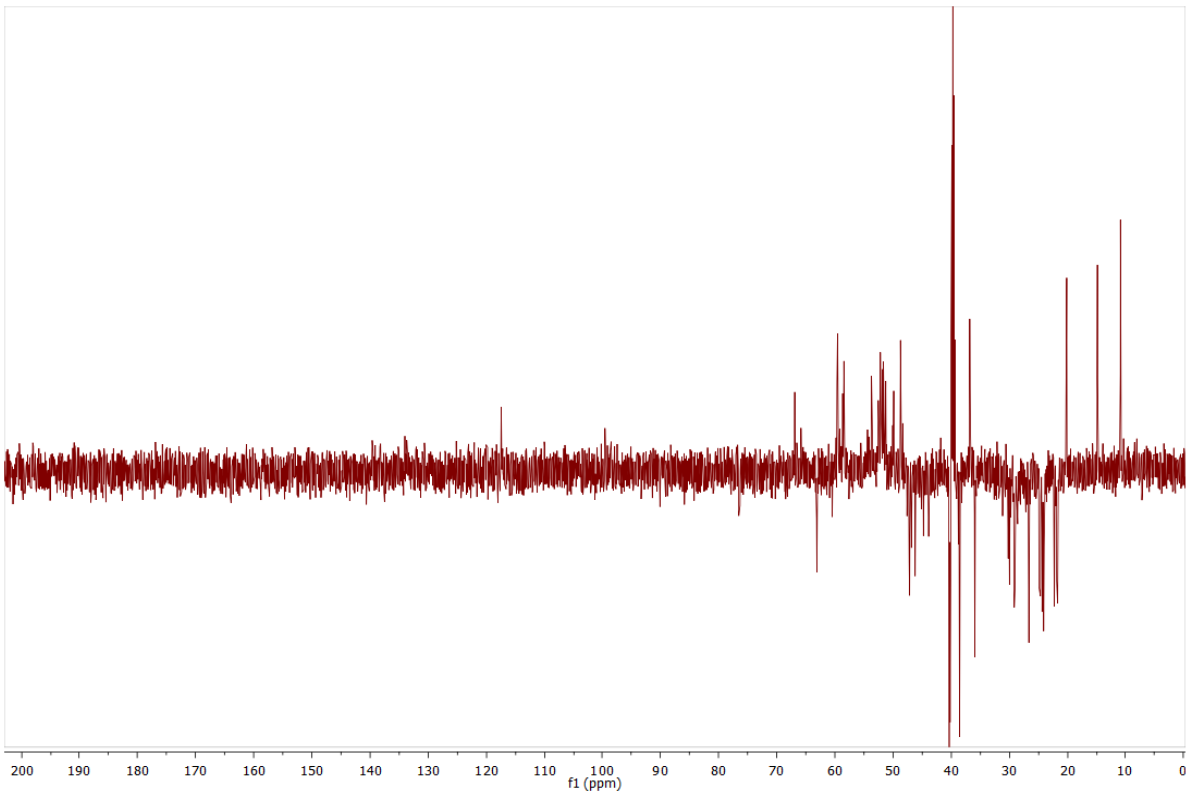


Figure S35. ^{13}C DEPT NMR spectrum of compound **43.**

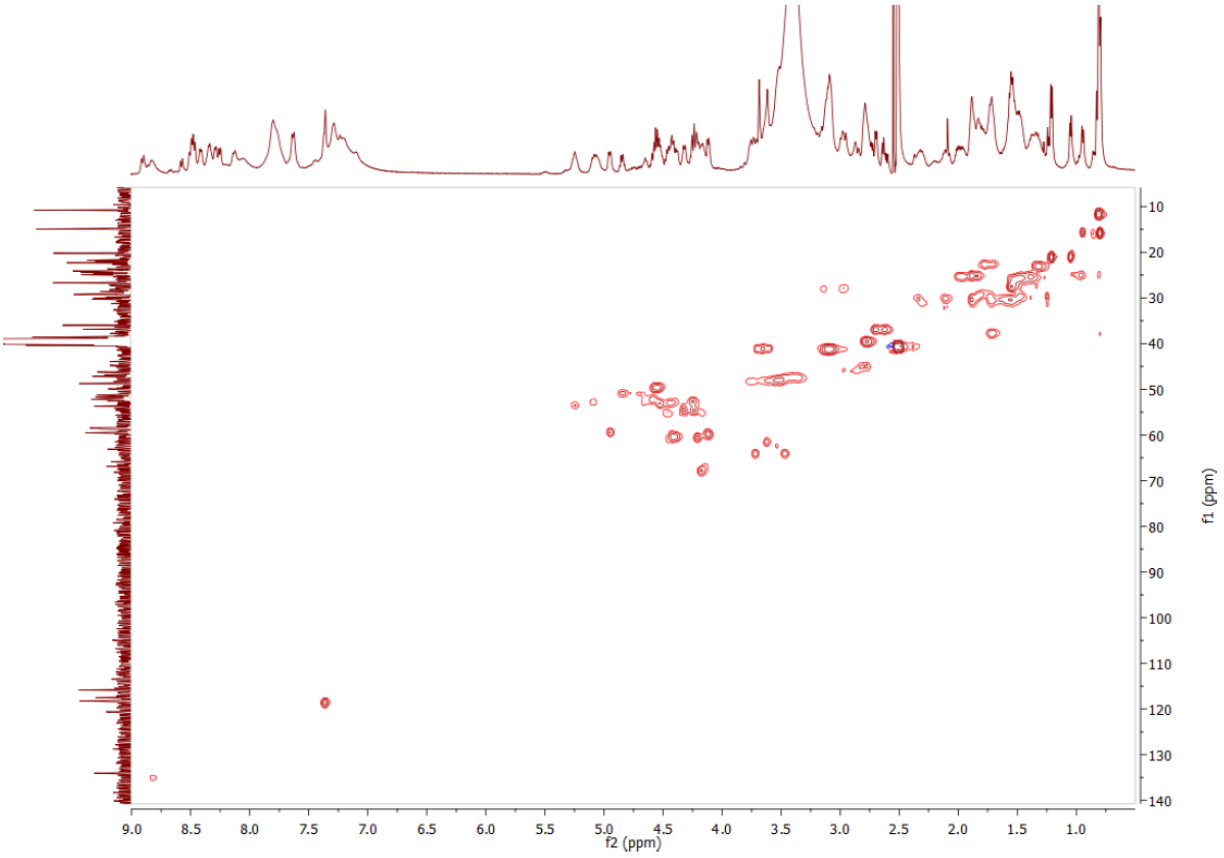


Figure S36. ^{1}H-^{13}C HSQC NMR spectrum of compound **43.**

Compound 44

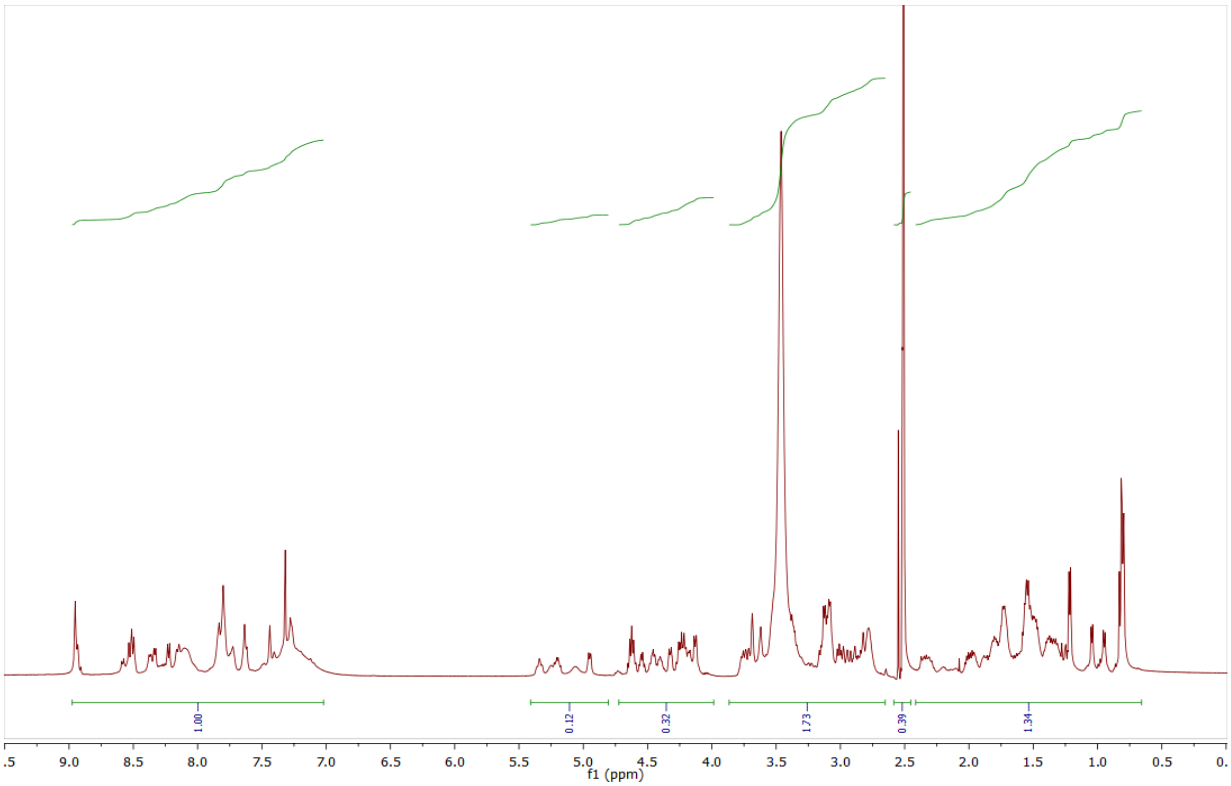


Figure S37. ^{1}H-NMR spectrum of compound **44.**

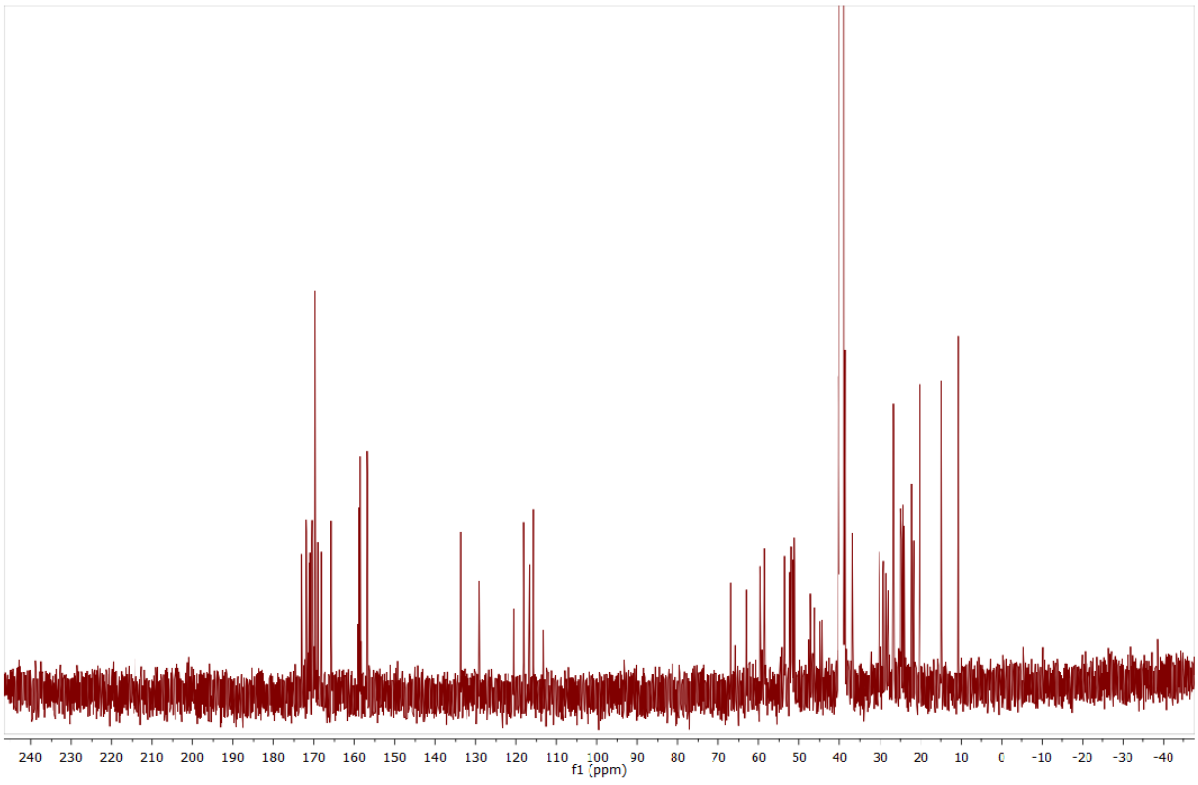


Figure S38. ^{13}C-NMR spectrum of compound **44.**

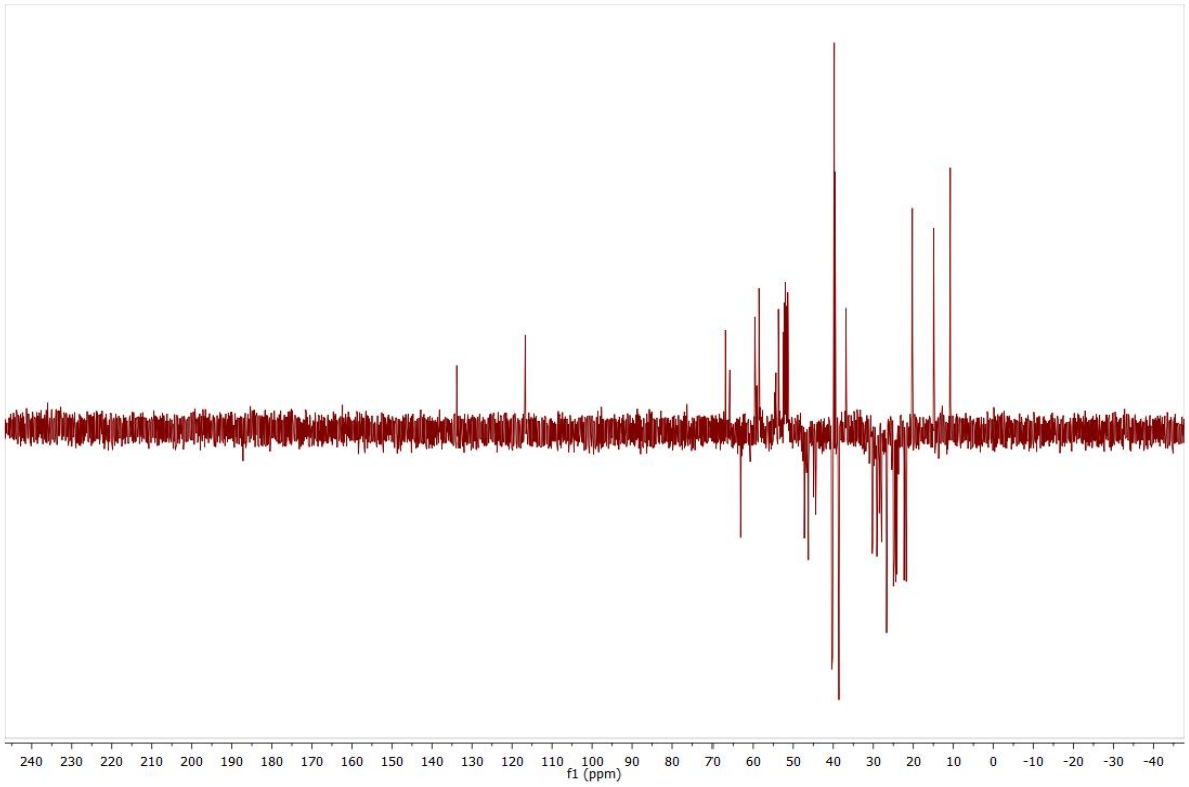


Figure S39. ^{13}C DEPT NMR spectrum of compound **44.**

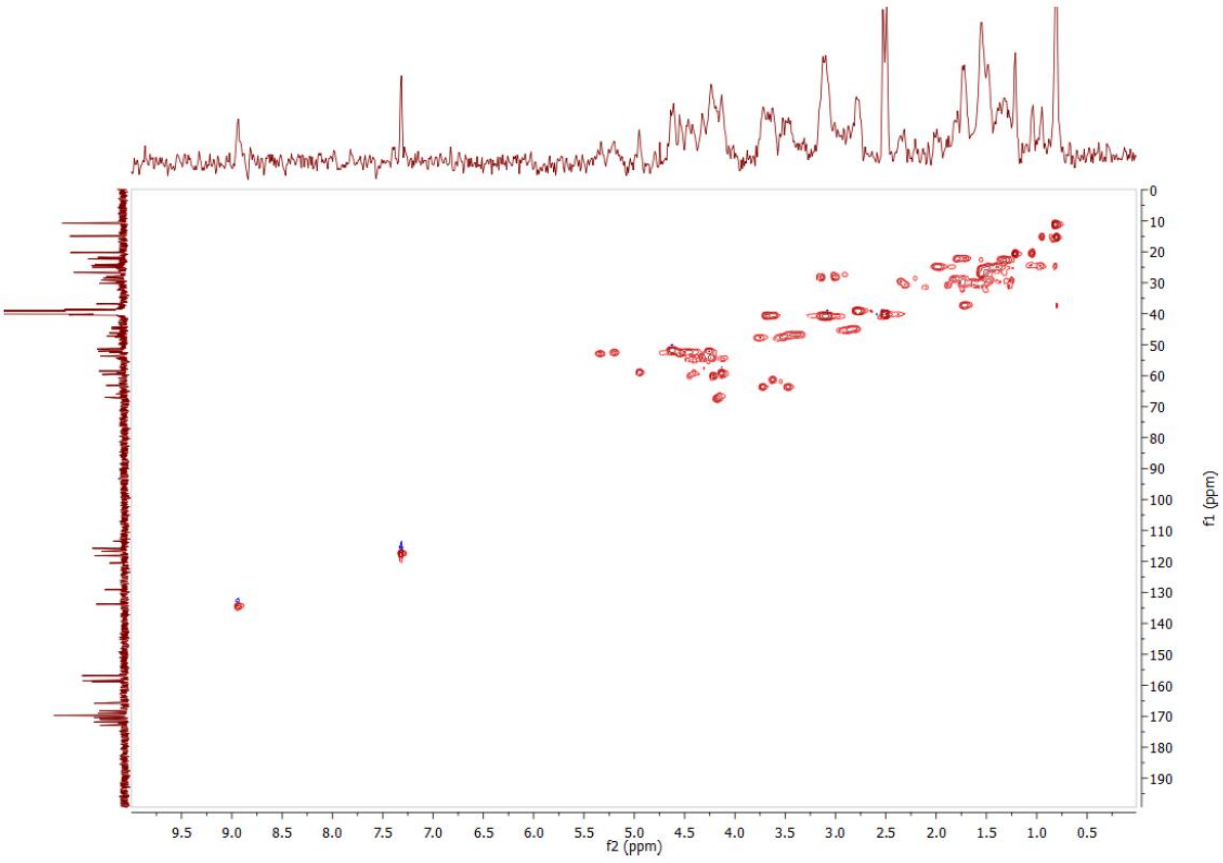


Figure S40. ^{1}H-^{13}C HSQC NMR spectrum of compound **44.**

4. Plotted Kinetic Data

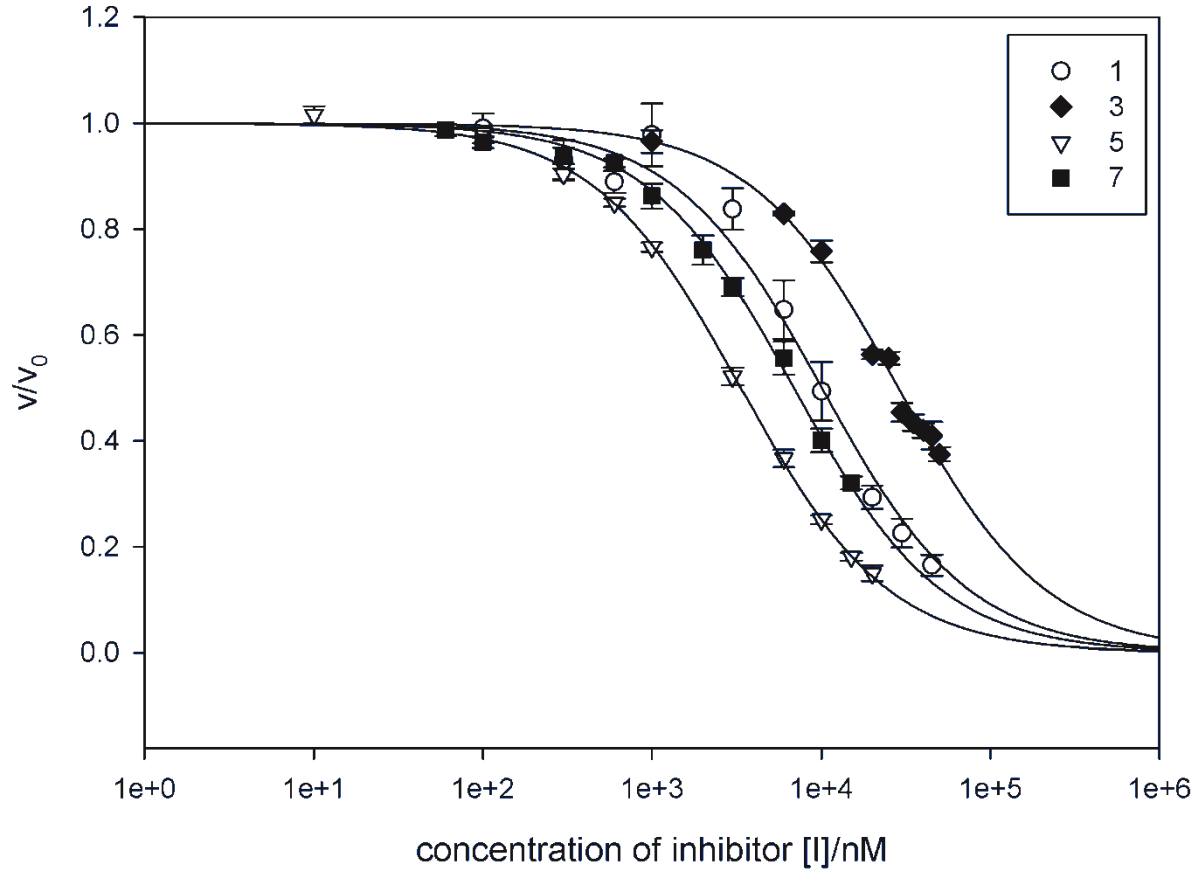


Figure S41. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of disulfide-bridged inhibitors **1** (white circle), **3** (black diamond), **5** (white traingles) and **7** (black squares). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

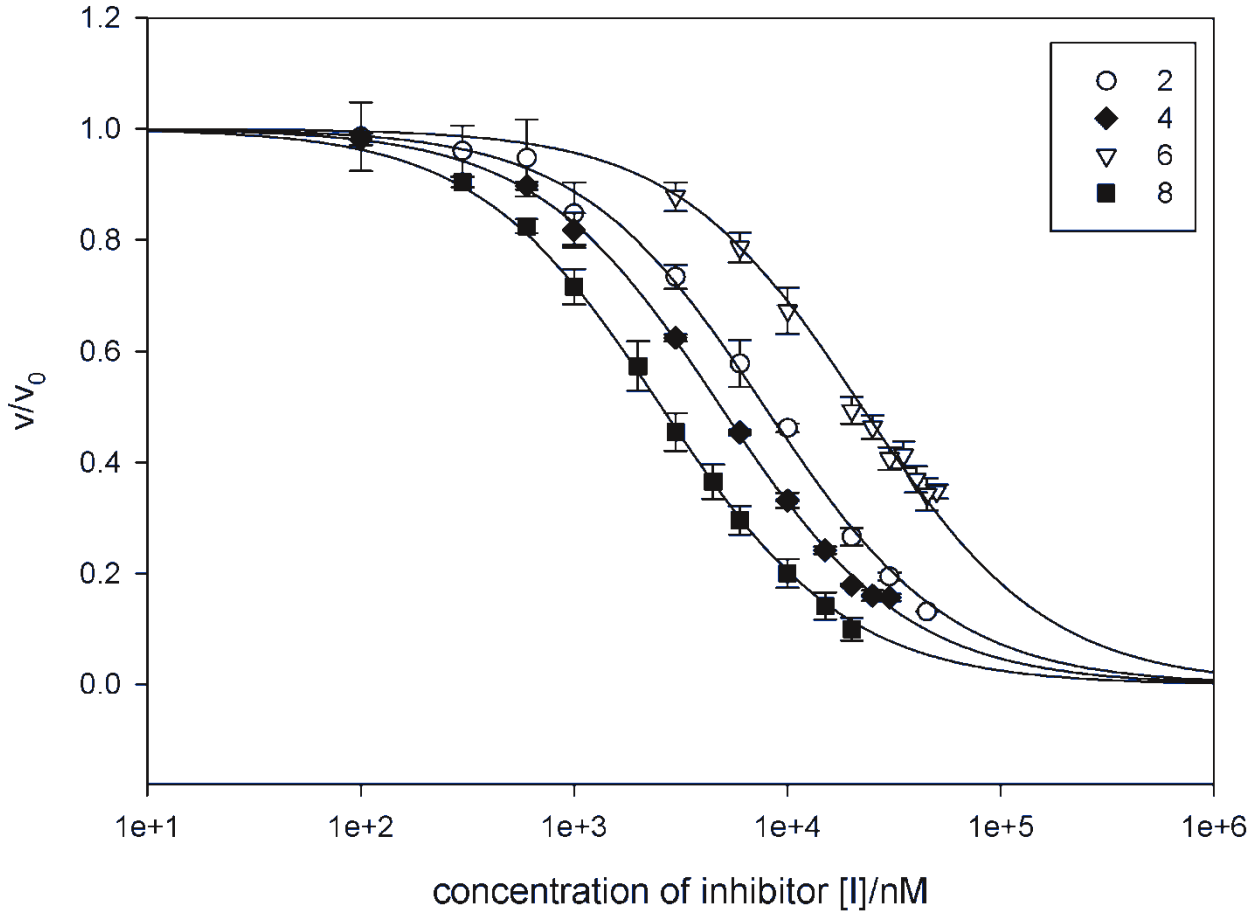


Figure S42. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of disulfide-bridged inhibitors **2** (white circle), **4** (black diamond), **6** (white traingles) and **8** (black squares). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

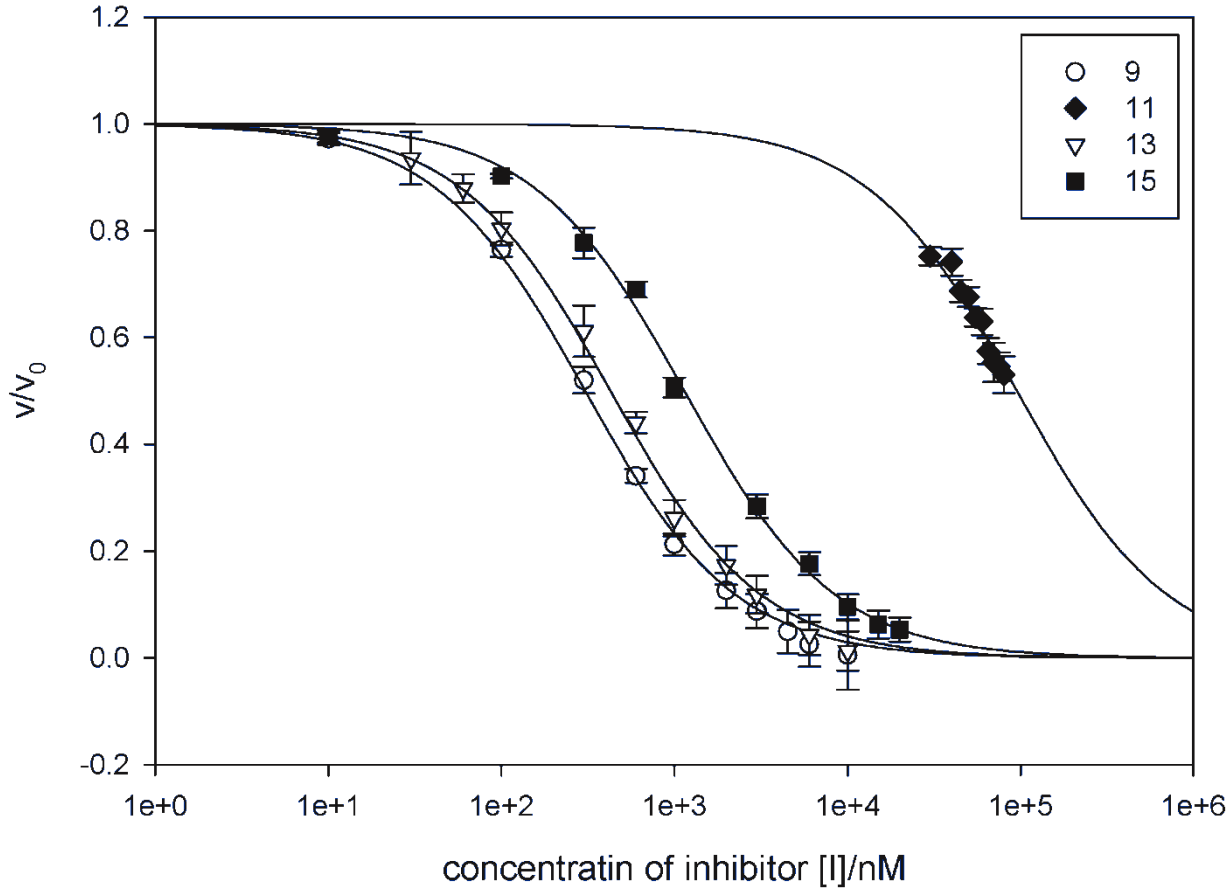


Figure S43. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of disulfide-bridged inhibitors **9** (white circle), **11** (black diamond), **13** (white traingles) and **15** (black squares). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

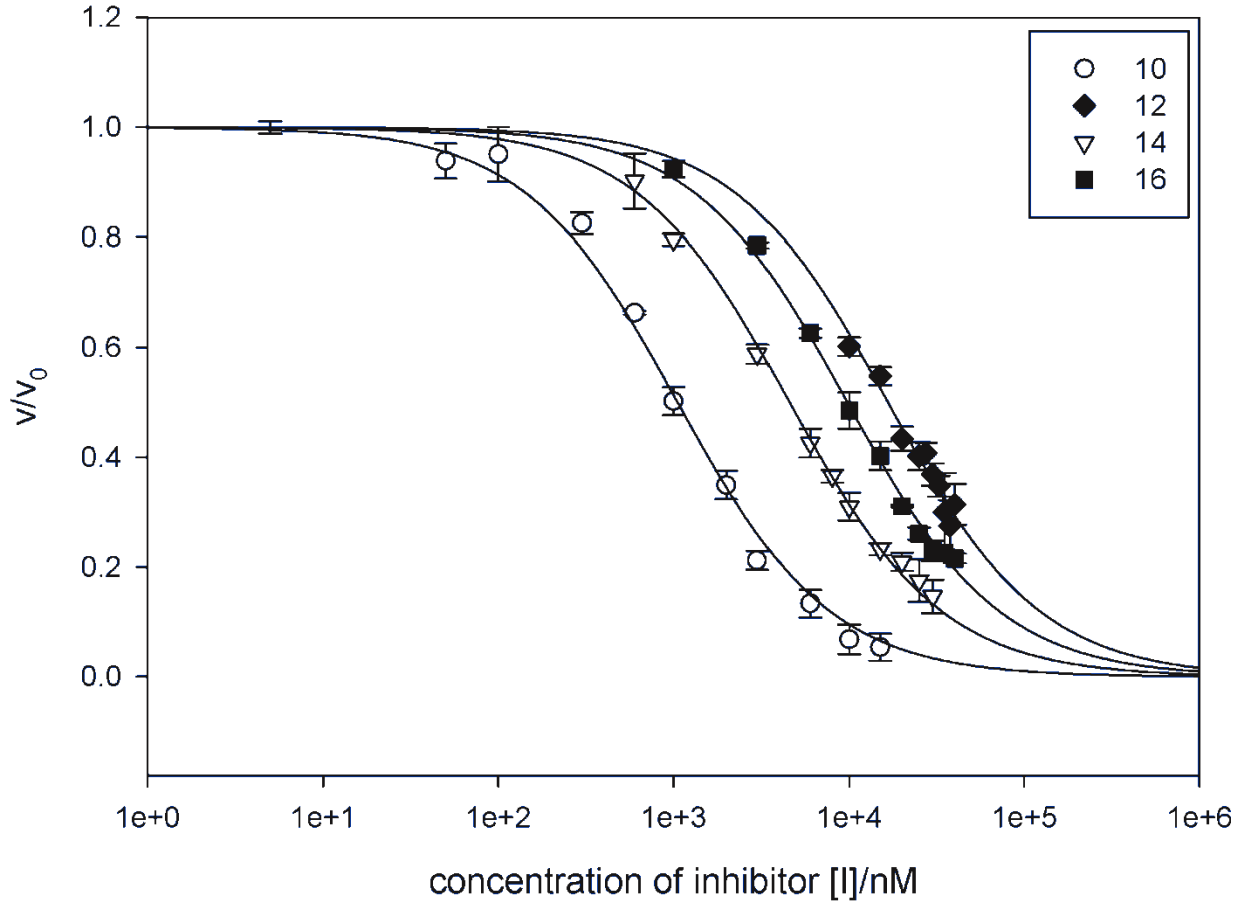


Figure S44. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of disulfide-bridged inhibitors **10** (white circle), **12** (black diamond), **14** (white traingles) and **16** (black squares). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

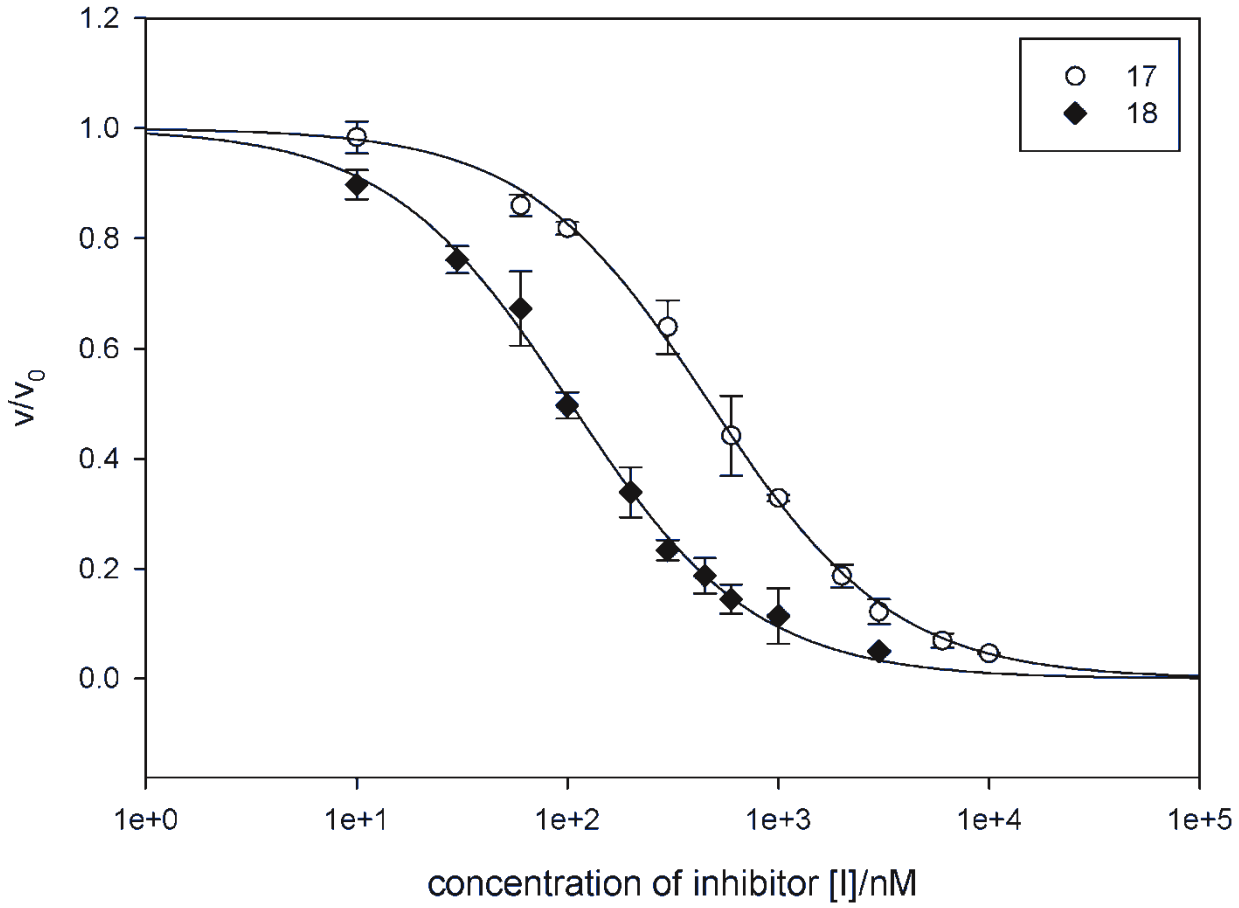


Figure S45. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of disulfide-bridged inhibitors **17** (white circle) and **18** (black diamond). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

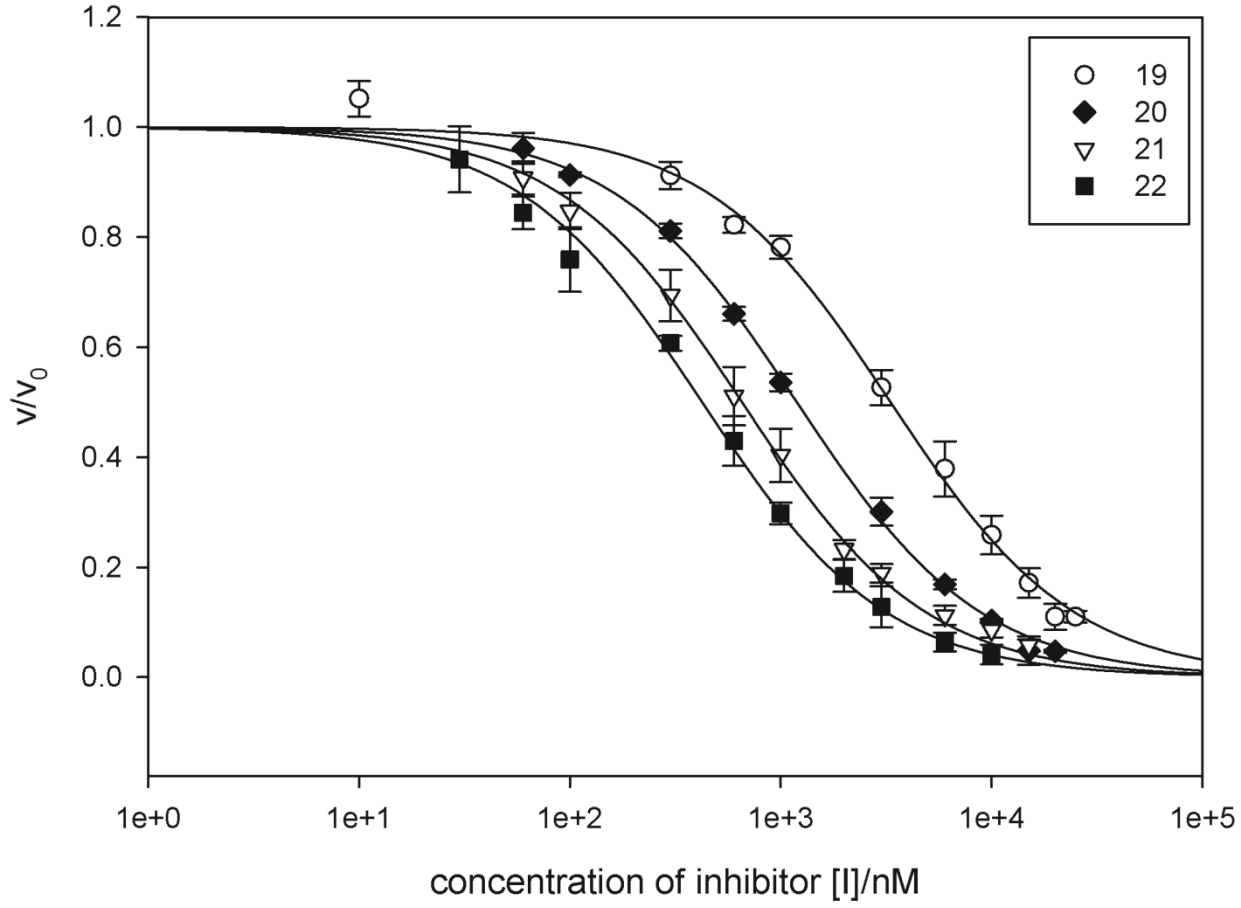


Figure S46. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of disulfide-bridged inhibitors **19** (white circle), **20** (black diamond), **21** (white traingles) and **22** (black squares). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

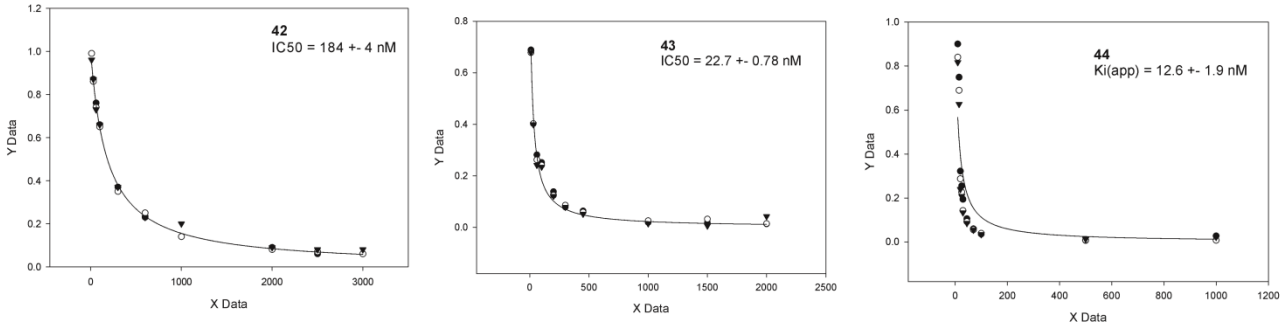


Figure S47. Dose-response curves (pH 7.6) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA of compound **42** (left), **43** (middle) and **44**(right).

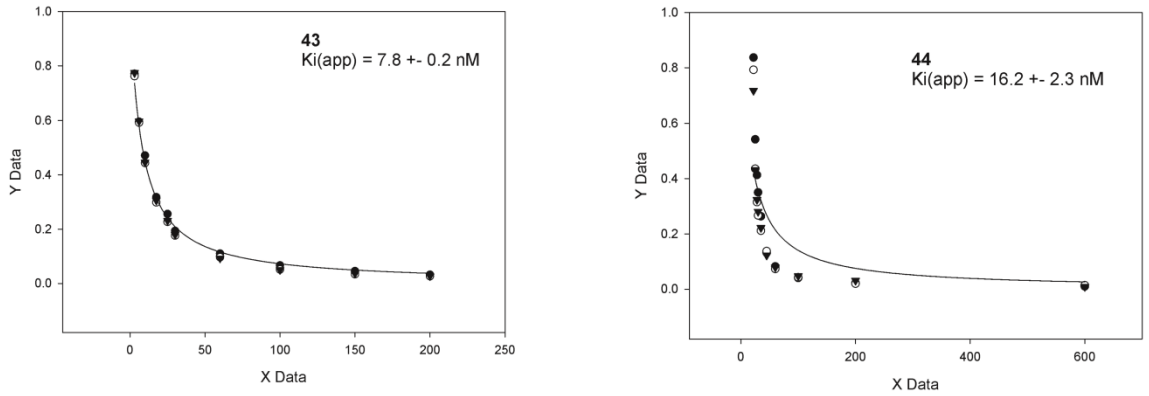


Figure S48. Dose-response curves (pH 8.5) for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA of compound **43** (left) and **44**(right).

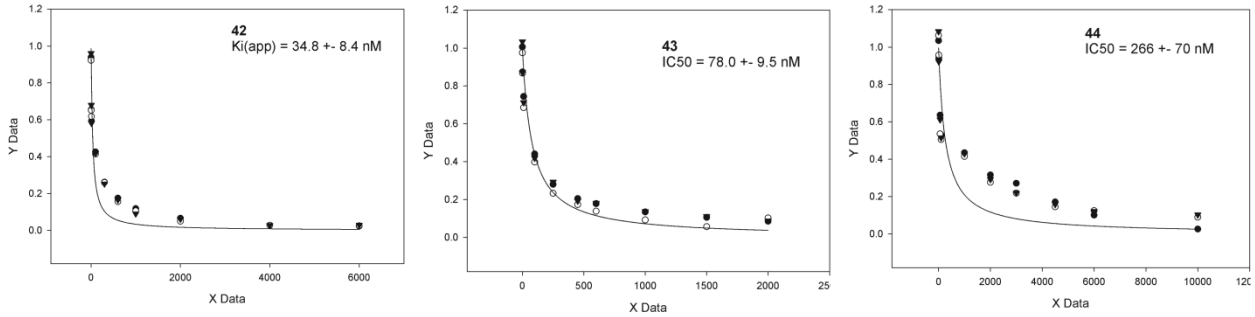


Figure S49. Dose-response curves (pH 7.6) for the inhibition of trypsin-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA of compound 42 (left), 43 (middle) and 44(right).

5. Additional Figures

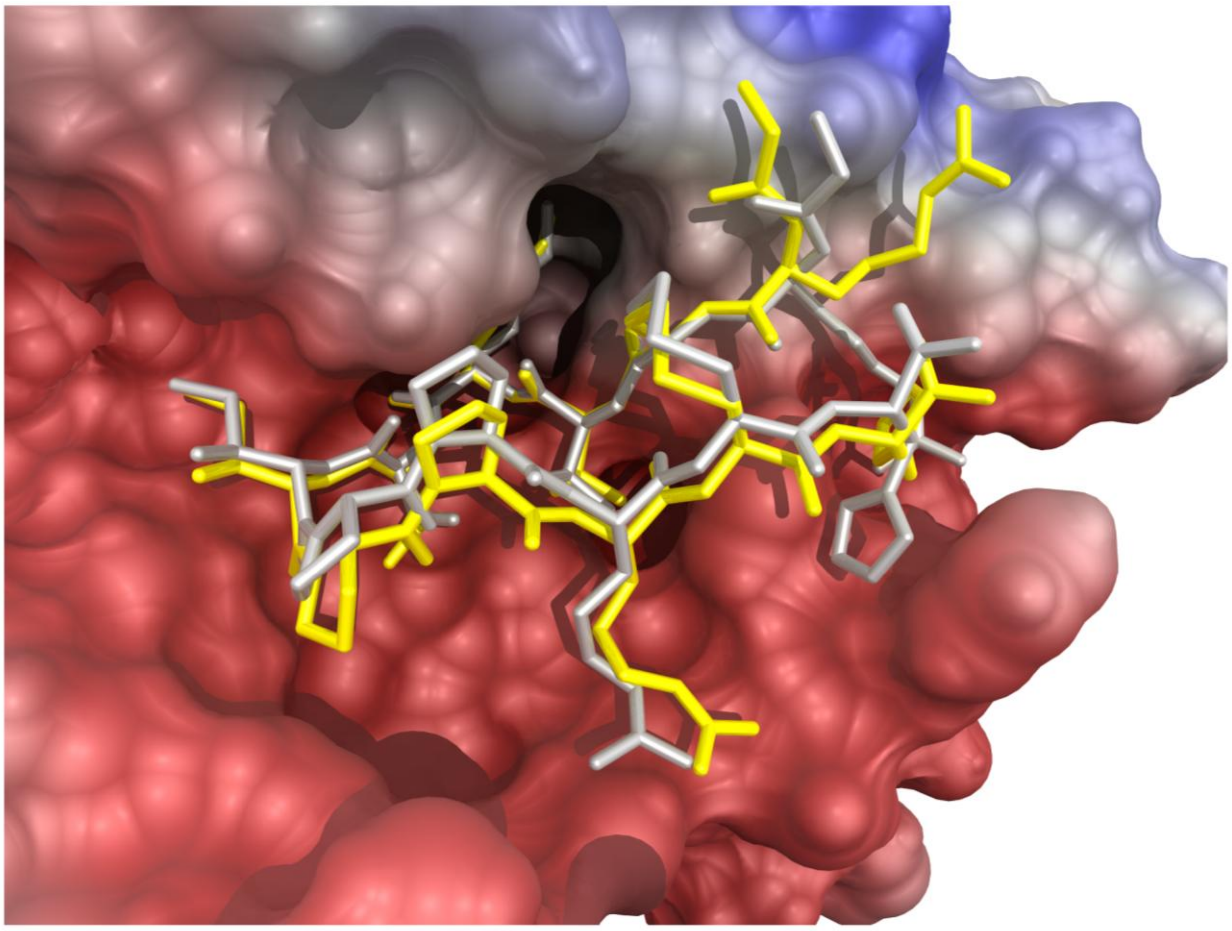


Figure S50. Comparison of SDMI-2-matriptase complexes generated by molecular replacement and energy minimization (white; see Figure 3 in the main text) and molecular dynamics simulations followed by docking (yellow; procedure was described in reference 22 from the main text).

6.3. Supporting Information for chapter 2.3

Contents: **Pages:**

ESI-MS **S2**

RP-HPLC **S4**

Plotted Kinetic Data **S6**

Table S1. ESI-MS data for compounds **3-11**.

Entry	SIM / g·mol^{-1}	$[M+H]^+$	$[M+2H]^{2+}$	$[M+3H]^{3+}$	$[M+4H]^{4+-}$
3	2025.3	-	1013.5	676.0	507.4
4	1635.9	-	818.9	546.2	409.9
5	2024.3	-	1013.5	676.0	-
6	1617.9	-	809.8	540.3	405.4
7	1564.8	1565.7	783.3	522.6	392.1
8	1617.9	1618.8	809.9	540.3	405.4
9	1502.8	-	752.3	501.9	376.6
10	1405.7	-	703.8	469.5	-
11	1422.7	-	712.4	475.2	-

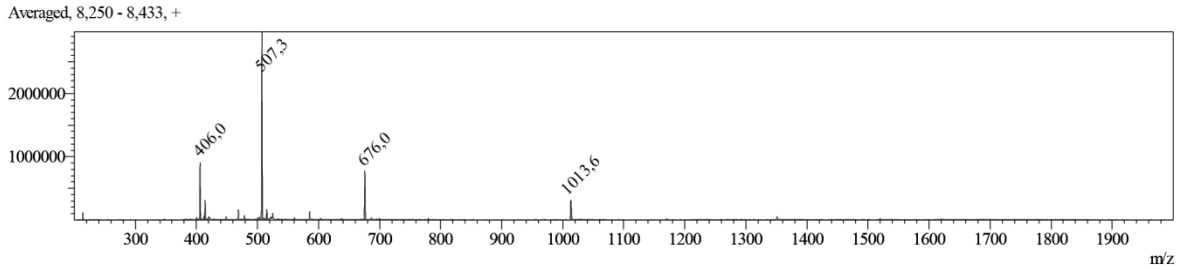


Figure S1. ESI-MS spectrum of compound **3** (positive polarization).

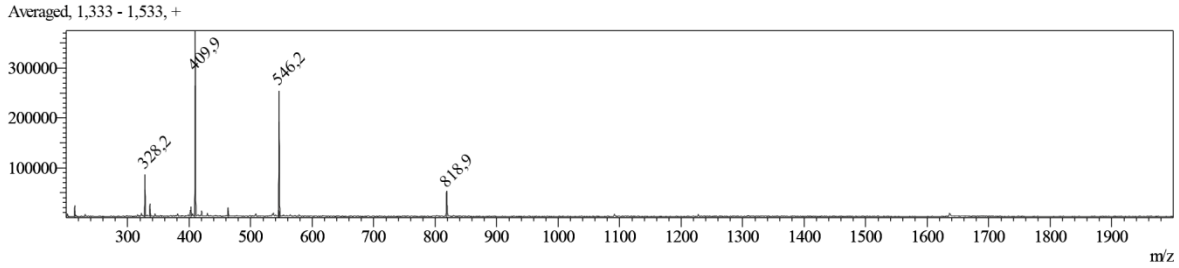


Figure S2. ESI-MS spectrum of compound **4** (positive polarization).

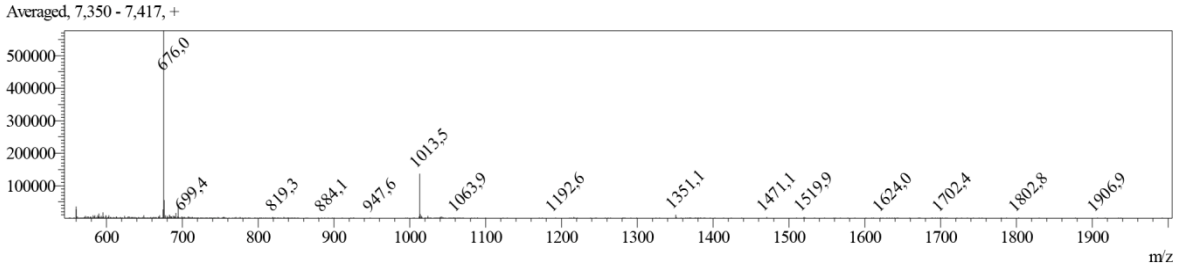


Figure S3. ESI-MS spectrum of compound **5** (positive polarization).

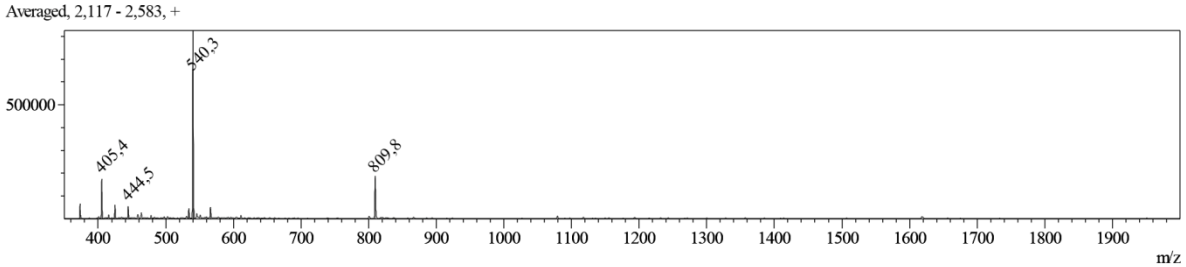


Figure S4. ESI-MS spectrum of compound **6** (positive polarization).

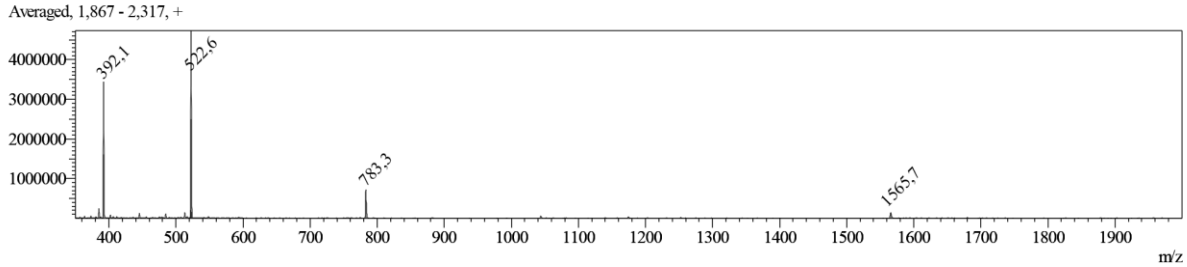


Figure S5. ESI-MS spectrum of compound **7** (positive polarization).

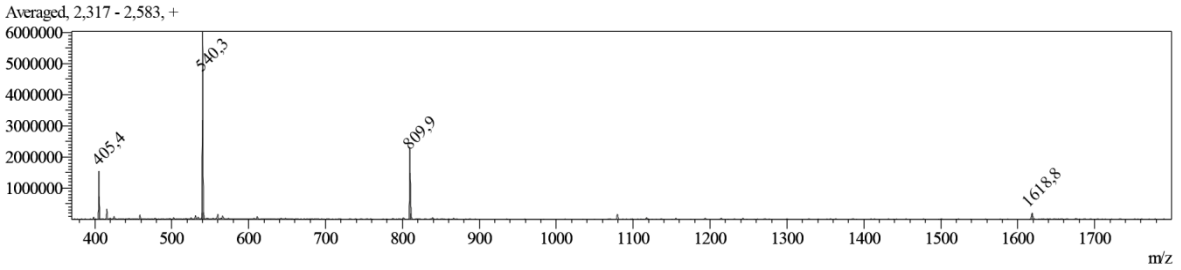


Figure S6. ESI-MS spectrum of compound **8** (positive polarization).

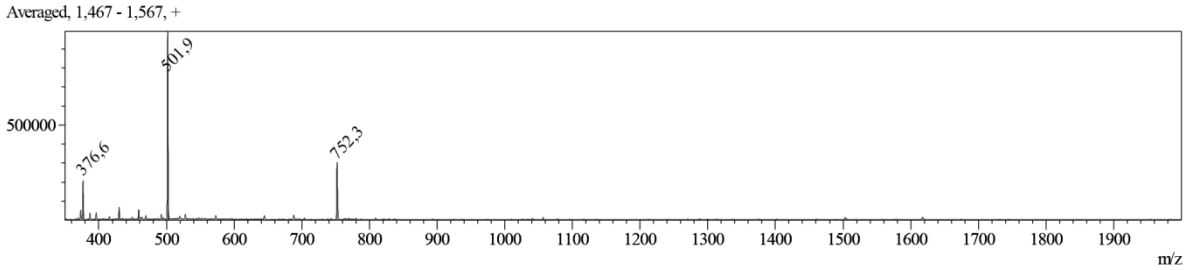


Figure S7. ESI-MS spectrum of compound **9** (positive polarization).

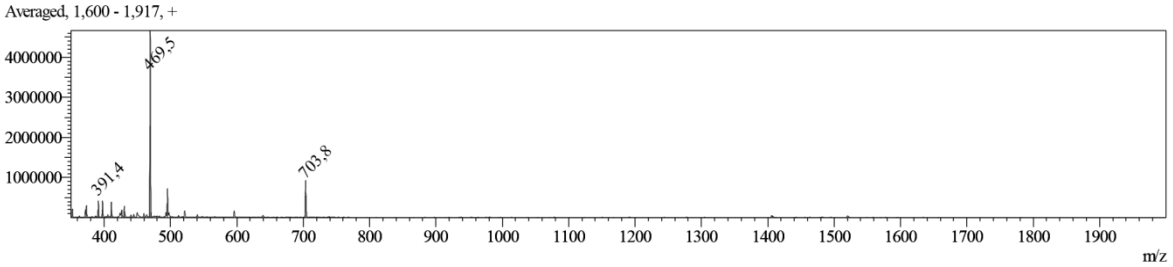


Figure S8. ESI-MS spectrum of compound **10** (positive polarization).

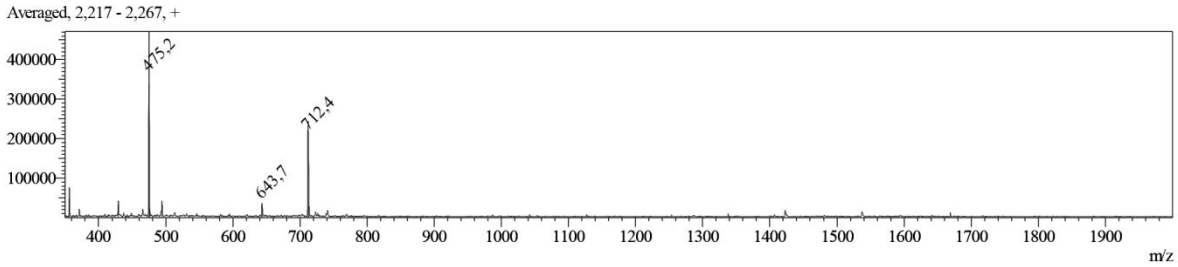


Figure S9. ESI-MS spectrum of compound **11** (positive polarization).

2. RP-HPLC Data

Table S2. Retention time of compounds **3-11** measured with RP-HPLC.

Entry	R_t /min	Gradient
3	22.0	9→54 %[a]
4	13.1	9→36 %[b]
5	22.4	9→45 %[c]
6	15.5	9→36 %[b]
7	17.4	9→36 %[b]
8	16.3	9→36 %[b]
9	16.1	9→36 %[b]
10	16.2	9→36 %[b]
11	14.7	9→36 %[b]

[a] 9 % acetonitrile over 2 min followed by 9→54 % acetonitrile in 0.1 % TFA over 20 min of flow rate 1 mL/min.
[b] 9 % acetonitrile over 2 min followed by 9→36 % acetonitrile in 0.1 % TFA over 20 min of flow rate 1 mL/min.
[c] 9 % acetonitrile over 2 min followed by 9→45 % acetonitrile in 0.1 % TFA over 20 min of flow rate 1 mL/min.

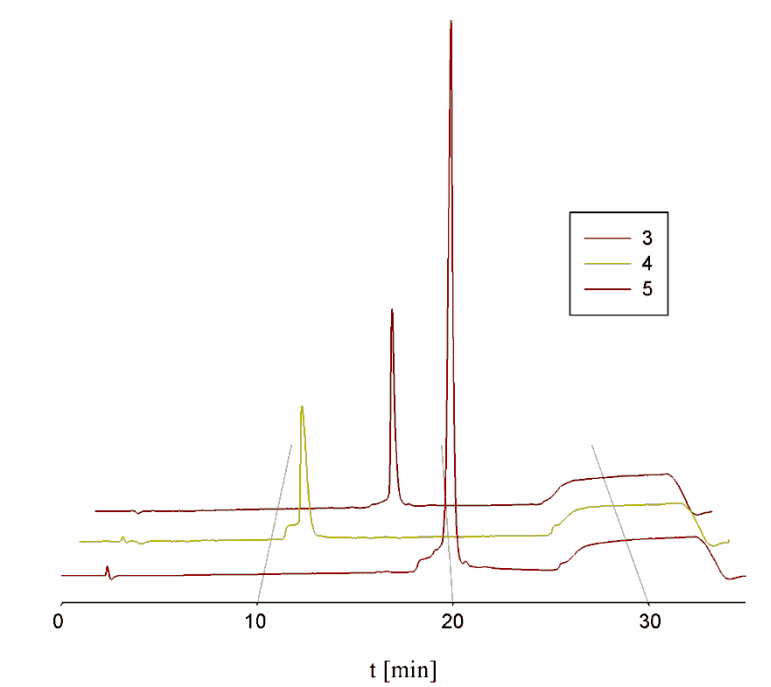


Figure S10. HPLC chromatogram of purified **3-5** recorded at 220 nm.

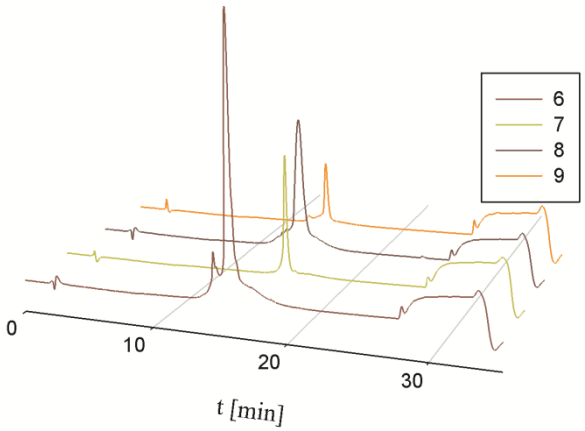


Figure S11. HPLC chromatogram of purified **6-9** recorded at 220 nm.

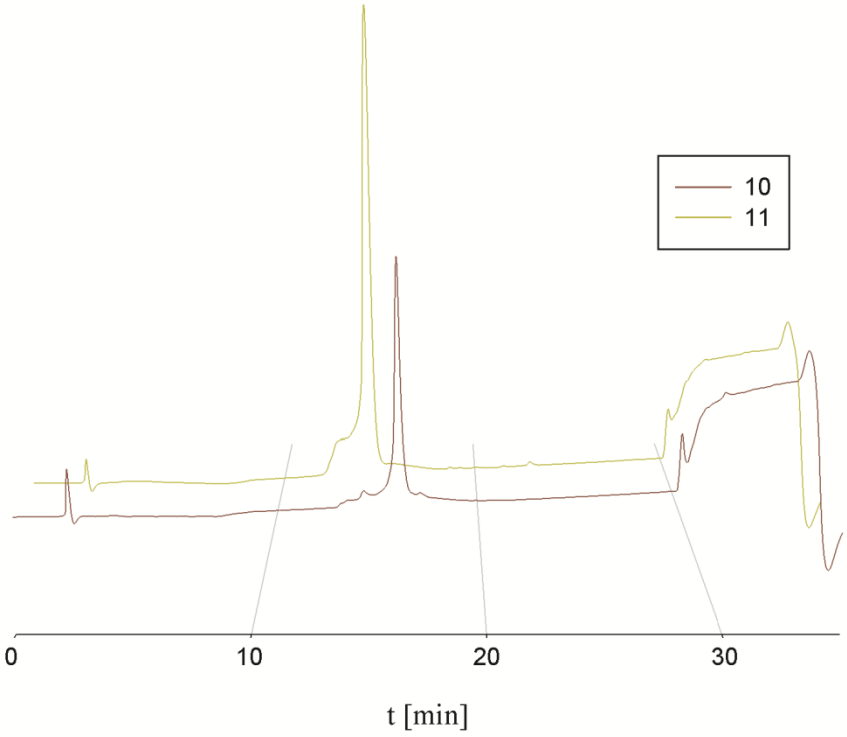


Figure S12. HPLC chromatogram of purified **10** and **11** recorded at 220 nm.

3. Plotted Kinetic Data

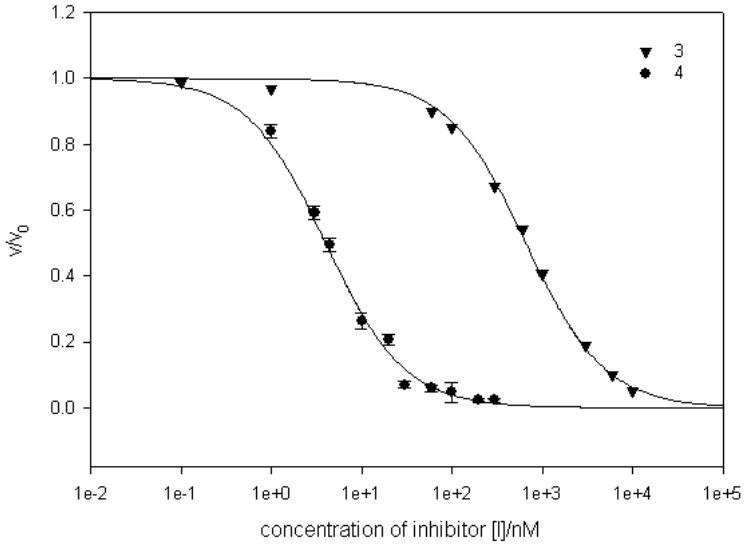


Figure S13. Dose-response curves for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA with the X-axis on a logarithmic scale. Comparison of matriptase inhibitors **3** (black triangles) and **4** (black circle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

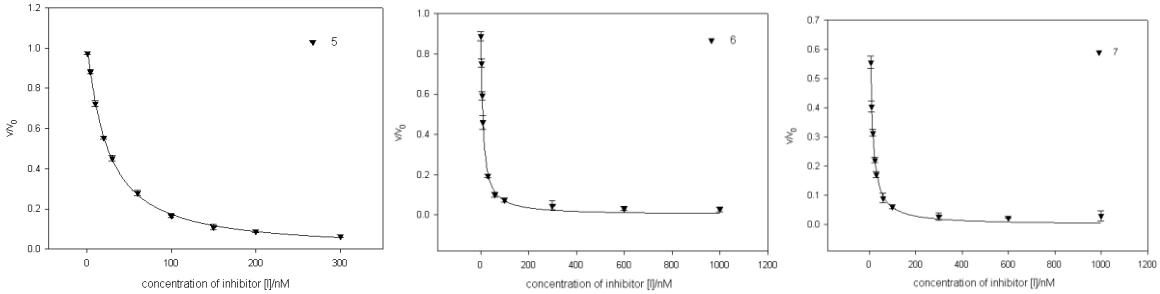


Figure S14. Dose-response curves for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA of compound **5** (left), **6** (middle) and **7** (right).

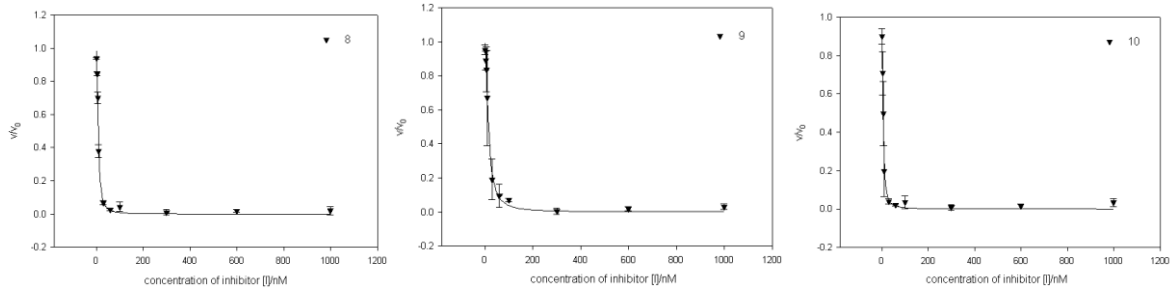


Figure S15. Dose-response curves for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA of compound **8** (left), **9** (middle) and **10** (right).

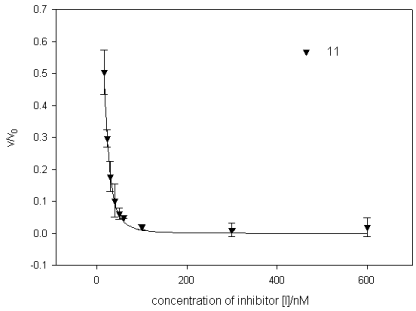


Figure S16. Dose-response curve for the inhibition of matriptase-catalyzed proteolysis of chromogenic substrate Boc-QAR-*p*NA of compound **11**.

6.4. Supporting Information for chapter 2.4

CHEMBIOCHEM

Supporting Information

Engineering a Constrained Peptidic Scaffold towards Potent and Selective Furin Inhibitors

Heiko Fittler,[a] Alexander Depp,[a] Olga Avrutina,[a] Sven O. Dahms,[b] Manuel E. Than,[b] Martin Empting,[c] and Harald Kolmar*[a]

cbic_201500447_sm_miscellaneous_information.pdf

Contents: **Pages**

Experimental Section	2
Modeling	4
ESI-MS	8
Plotted Kinetic Data	12
HPLC	17
Literature	20

Experimental section

Peptide Synthesis

Peptides **1**-**10**, **15** were synthesized using Fmoc-SPPS on a Tenta-Gel S AC resin (Rapp Polymere) preloaded with Fmoc-L-Asp(tBu). Peptides **11**-**14**, **16**-**17** were assembled on an AmphiSpheres 40 RAM resin (Agilent).

All compounds were synthesized on a CEM Liberty Blue system at a 0.05 mmol scale using double couplings with 5 eq of each component, 0.2 M of the corresponding amino acid, 0.5 M DIC, and 1 M Oxyma in DMF. All amino acids except for cysteine, histidine, and arginine were coupled by a two step program: 1) 75°C, 170W, 15s 2) 90°C, 50W, 110 s. Histidine and cysteine were coupled as follows: 1) 25°C, 0W, 3 min 2) 50°C, 35W, 5min. Arginine was incubated 30 min at room temperature: 1) 25°C, 0W, 30min 2) 75°C, 30W, 2min. Fmoc deprotection with 20% piperidine (v/v) and 0.1 M oxyma in DMF was conducted as following: 75°C, 100W, 3min.

On-support copper(I)-catalyzed azide-alkyne cylcoaddition (CuAAC) of peptide **16** was performed on a 0.05 mmol scale using 1 eq. copper (II) sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$), 2 eq. sodium ascorbate (NaAsc) and 4 eq. *N*,*N*-diisopropylethylamine (DIEA) in 6 mL argon-flushed DMF overnight at ambient temperature. Then the solution was filtrated and the resin was washed with methanol (3×5mL), 0.5% sodium diethyldithiocarbamate in DMF (w/v, 3×5mL), DMF (3×5mL), and DCM (3×5mL).

On-support ruthenium(II)-catalyzed azide-alkyne cycloaddition (RuAAC) for peptide **17** was conducted on a 0.05 mmol scale keeping the *N*-terminal Fmoc-protection intact and using 3.8 mg Cp*RuCl(COD) (20 mol% of initial loading of the resin). Argon was bubbled through the suspension of the resin in DMF. Then the catalyst was added; the syringe was sealed and placed in the manual Discover SPS microwave peptide synthesizer with the fiber-optic temperature probe measuring the temperature of the reference. The RuAAC was conducted using the following program: 60°C, 30 W, 5h. Afterwards, the resin was washed as described for the CuAAC procedure. Finally, the Fmoc group was deproteced using 20% piperidine in DMF (v/v) for 5 minutes at ambient temperature.

After drying *in vacuo*, all peptides were cleaved from the solid support using a standard cleavage cocktail: trifluoroacetic acid (TFA)/water/anisole/triethylsilane (47:1:1:1, v:v:v:v) and dithiothreitol to suppress oxidation of cysteines if necessary. The peptides were precipitated after 2 h in 40 mL cold diethyl ether and washed twice with diethyl ether (2×40 mL).

Macrocylization of peptides **1**-**15** *via* formation of an intramolecular disulfide was achieved upon air oxidation during gentle agitation of diluted solutions in 100 mM $(NH_4)_2CO_3$ aq. (1 mg mL^{-1}, pH= 8.4) overnight. The folding mixture was afterwards removed by freeze-drying and the peptides were isolated by semi-preparative RP-HPLC and the mass of the lyophilized peptides were determined by weighing. The concentrations were prepared from a 1 mM stock solution in water.

Enzyme Production

Matriptase-1 was recombinantly produced as previously described.[1]

Human Furin was prepared as described before.[2] In short, the protein was expressed by transient transfection of human embryonic kidney cells and possessed the catalytic and P-domains. It was purified by a three-step procedure comprising metal affinity chromatography, inhibitor-based affinity chromatography, and size exclusion chromatography.[3]

Enzymatic assays

Furin

The inhibition constants for peptide **1**-**17** were determined in triplicate by incubating the peptides at different concentrations for 10 minutes with human furin (final concentration: [E]=0.95 nM) in the furin inhibition buffer (100 mM HEPES buffer, pH 7.0 containing 1mg/mL BSA, 0.02% NaN_3 (w/v), 2 mM $CaCl_2$ and 0.2% TritonX·1000 (v/v)). Then the substrate Boc-Arg-Val-Arg-Arg-AMC (final concentration: [S]= 50 µM, K_M=36 µM[4]) was added and the measurement with an infinite 200Pro TECAN Genius plate reader followed. The dose-response curves were fitted with Eqn (1) to yield the IC50 value.

Matriptase and Trypsin

The inhibition constants for peptide **11**, **12**, **16**, and **17** were determined in triplicate by incubating the peptides at different concentrations for 10 minutes with matriptase-1 or trypsin (final concenctration: [E]=1 nM) in the inhibition buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.4). Then the chromogenic substrate Boc-Gln-Ala-Arg-*p*NA (final concentration: [S]= 250 µM) was added and the measurement with an infinite 200Pro TECAN Genius plate reader followed. The dose-response curves were fitted with Eqn (1) to yield the IC50 value.[5-6]

$$\frac{v}{v_0} = \left(1 + \left(\frac{[I]}{IC_{50}}\right)^{Hill}\right)^{-1} \quad (1)$$

The substrate-independent inhibition constant K_i was obtained using Eqn (2).

$$K_i = IC_{50}\left(1 + \frac{[S]}{K_m}\right)^{-1} \quad (2)$$

General

All chemicals and solvents used in this study were delivered by Iris Biotech (Marktredwitz, Germany), Bachem (Bubendorf, Switzerland), Orpegen (Heidelberg, Germany), Novabiochem (Darmstadt, Germany), Sigma-Aldrich (Steinheim, Germany), Roth (Karlsruhe, Germany), and Agilent (Böblingen, Germany).

Peptides were analyzed using an analytical RP-HPLC from Varian, model 920-LC applying a Phenomenex Hypersil 5u BDS C18 LC column (150×4.6mm, 5µm, 13 nm), and isolated by semi-preparative RP-HPLC (Varian) using a Phenomenex Luna 5u C18 LC column (250×12.2mm, 5µm, 10 nm) as stationary phase. Both eluents A (water) and B (90% aq. acetonitrile) contained 0.1% aq. TFA.

Electrospray ionization mass spectra were collected on a Shimadzu LCMS-2020 equipped with a Phenomenex Jupiter 5u C4 LC column (50×1mm, 5µm, 30nm) using an eluent system of 0.1% aq. formic acid, LC-MS grade (eluent A) and acetonitrile containing 0.1% aq. formic acid, LC-MS grade (eluent B).

Molecular modeling

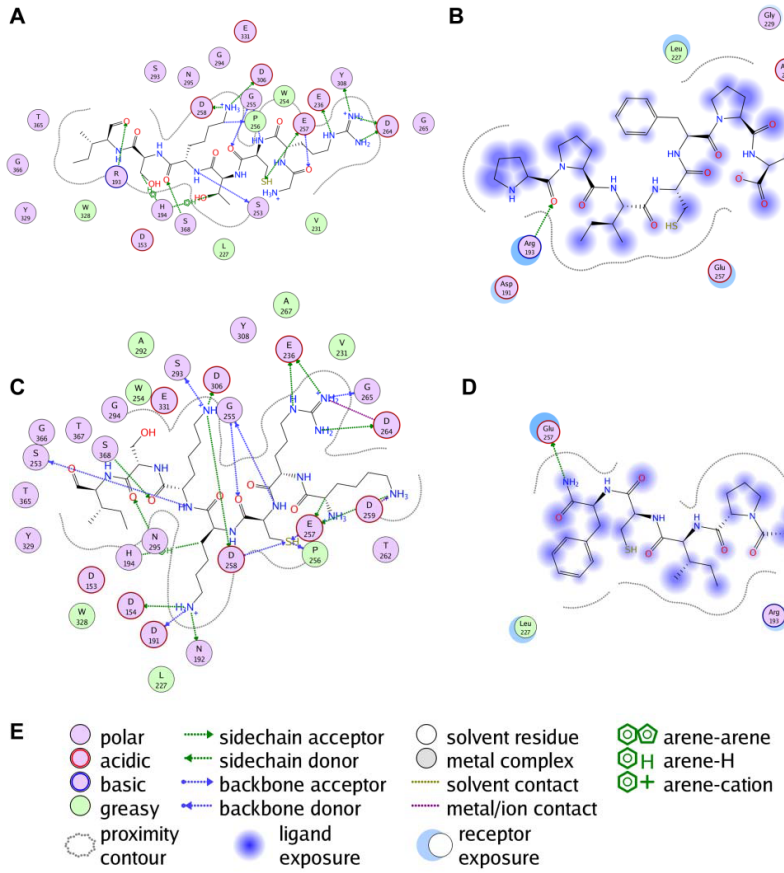


Figure S1: 2D depiction of interactions of residues 1-7 of compounds **1** (A) and **11** (C) as well as thier C-terminal sections (B, D) as provided by the "Ligand interactions"-command of Molecular Operating Environment (MOE, Chemical Computing Group). Water molecules, and ions were omitted for clarity. E) Legend.

Structure of the furin-inhibitor complexes were generated based on published X-ray structures (furin and covalent inhibitor: 1P8J; matriptase-1 and SFTI-1: 3P8F) by running a self-written macro with YASARA structure (see below) using the YASARA2 force field.[7] In brief, we aligned both proteases based on the catalytic triad and refined this superimposition by aligning the ligand backbones. Then, we made the respective side chain substitutions and optimized side chain geometries with the built-in "optimize" command. Finally, we performed a four-step energy minimization: first, with fixed protein and ligand backbone, then, with fully flexible ligand followed by freeing the protein side chains and a final fully unconstrained minimization step. Ligand interactions were analysed and visualized with Molecular Operating Environment (MOE, chemical computing group)[7] and 3D images were rendered with POVray.

Macro:

```
Clear
ShowMessage "Save 1P8J.pdb and 3P8F.pdb (from www.pdb.org) in your ...\YASARA\pdb folder!"
#Wait Continuebutton

ShowMessage "Processing Furin..."
LoadPDB 1p8j
DelMol B C D E F G H I K L M N O P Q R S T U V W X Y Z
```

```
DelMol Atom 3641
DelMol Atom 3641
DelMol Atom 3641
DelMol Atom 3641
DelMol Atom 3641

ShowMessage "Processing Matriptase..."
LoadPDB 3p8f
DelMol Atom 5610
DelMol Atom 5610

ShowMessage "Aligning Matriptase to Furin..."
#This is an initial alignment via the catalytical triad:
AlignAtom 3939-3948 4410-4417 5118-5123,356-363 681-690 1961-
1966,MatchElement=Yes,MatchName=No,MatchBonds=Yes,MatchBondOrders=Yes,MatchSecStr=No,DisMax=0.6
#This is an alignment via the ligand
AlignMol I,J,Method=MUSTANGPP

CenterObj 1
OriAll Alpha=3.6e+001,Beta=3.5e+002,Gamma=1.5e+001
AutoMoveToAll X=7, Y=20, Z=10,Steps=10

ShowHUDMol Atom 1-3640
StickMol Atom 3592
ShowHUDMol Atom 3641-5609
StickMol Atom 5505
ColorMol Atom 3592,Red

ShowMessage "Modelling SFTI-1-derived Furin Inhibitor (SDFI)..."
DuplicateAll
RemoveObj 1
RemoveObj 2
DelMol Atom 3592
DelMol Atom 3592
NameObj 3,Furin
NameObj 4,SFTI-1

FixObj 3
DuplicateObj 4
NameObj 5,SDFI
RemoveObj 4
ShowHUDRes Atom 3592-3696
SwapRes Atom 3592,Lys,Isomer=L
SwapRes Atom 3618,Lys,Isomer=L
DelRes Atom 3696
DelRes Atom 3689
CleanAll
StickObj 5
SwapAtom Atom 7199,Nitrogen,UpdateBonds=Yes,UpdateHyd=Yes
SwapBond 7197,7198,2.00,Update=Yes
SwapBond 7197,7199,1.00,Update=Yes
HideAtom Element H

ShowMessage "Press CONTINUE to perform optimization of inhibitor side chains."
#Wait Continuebutton
ShowMessage "Performing optimization of inhibitor side chains..."
FixRes Ser Cys Pro
Cell Auto, Extension=10
SimSpeed Normal
ForceField YASARA2,SetPar=Yes
Interactions Bond,Angle,Dihedral,Planarity,Coulomb,VdW
OptimizeObj 5,Method=SCWALL,Structures=1
FreeObj 5

ShowMessage "Press CONTINUE to perform four-step energy minimization (in vacuo)."
#Wait Continuebutton
ShowMessage "Preparing scene..."
```

```
DuplicateObj 5
NameObj 7,SDFI_Emin
RemoveObj 5
FixAtom backbone
Experiment Minimization
Experiment On
Wait ExpEnd

FreeAtom All
FixObj 3
Experiment Minimization
Experiment On
ShowHBoObj 7,Extend=Yes
Wait ExpEnd

DuplicateObj 3
RemoveObj 8
FreeAtom All
FixAtom backbone
FreeObj 7
Experiment Minimization
Experiment On
ShowHBoObj 7,Extend=Yes
Wait ExpEnd

FreeAtom All
Experiment Minimization
Experiment On
ShowHBoObj 7,Extend=Yes
Wait ExpEnd

AddObj 1
HideMol Atom 1
AddObj 5

ShowMessage "Molecular replacement done. Creating SFTI-1[1-14]-Furin model."
#Wait Continuebutton
ShowMessage "Preparing scene..."
AddObj 4
AddObj 8
#DuplicateObj 3
DuplicateObj 4
RemoveObj 1
RemoveObj 3
RemoveObj 4
RemoveObj 5
RemoveObj 7
DelBond 7004,7103
CleanObj 9
StickObj 9
HideAtom Element H
NameObj 9,oSFTI

ShowMessage "Press CONTINUE to perform optimization of SFTI-1[1-14] side chains."
#Wait Continuebutton
ShowMessage "Performing optimization of SFTI-1[1-14] side chains..."
FixRes Ser Cys Pro
Cell Auto, Extension=10
OptimizeObj 9,Method=SCWALL,Structures=1
FreeObj 9

ShowMessage "Press CONTINUE to perform four-step energy minimization (in vacuo)."
#Wait Continuebutton
ShowMessage "Preparing scene..."
DuplicateObj 9
NameObj 10,oSFTI_Emin
RemoveObj 9
```

```
FixAtom backbone
Experiment Minimization
Experiment On
Wait ExpEnd
FreeAtom All
FixObj 8
Experiment Minimization
Experiment On
ShowHBoObj 10,Extend=Yes
Wait ExpEnd

FreeAtom All
FixAtom backbone
FreeObj 10
Experiment Minimization
Experiment On
ShowHBoObj 10,Extend=Yes
Wait ExpEnd

FreeAtom All
Experiment Minimization
Experiment On
ShowHBoObj 10,Extend=Yes
Wait ExpEnd

ShowMessage "Modelling done. Creating SDFI-AhaPra model."
#Wait Continuebutton
ShowMessage "Preparing scene..."
AddObj 3
AddObj 7
DuplicateObj 3
DuplicateObj 7
RemoveObj 3
RemoveObj 7
RemoveObj 8
RemoveObj 10
NameObj 12,SDFI_Aza-Pra
SwapAtom Atom 7192,Carbon,UpdateBonds=Yes,UpdateHyd=Yes
SwapAtom Atom 7061,Carbon,UpdateBonds=Yes,UpdateHyd=Yes
DelBond 7061,7194
CleanObj 12
SwapAtom Atom 7196,Carbon,UpdateBonds=Yes,UpdateHyd=Yes
SwapBond 7195,7198,2.00,Update=Yes
AddBond 7061,7197,1.00,Update=Yes
SwapAtom Atom 7196,Nitrogen,UpdateBonds=Yes,UpdateHyd=Yes
CleanObj 12
SwapAtom Atom 7197,Nitrogen,UpdateBonds=Yes,UpdateHyd=Yes
SwapAtom Atom 7195,Carbon,UpdateBonds=Yes,UpdateHyd=Yes
SwapAtom Atom 7196,Nitrogen,UpdateBonds=Yes,UpdateHyd=Yes
AddBond 7198,7203,2.00,Update=Yes
SwapBond 7194,7199,1.00,Update=Yes
SwapBond 7195,7194,2.00,Update=Yes

ShowMessage "Press CONTINUE to perform four-step energy minimization (in vacuo)."
#Wait Continuebutton
ShowMessage "Preparing scene..."
FixAtom backbone
Experiment Minimization
Experiment On
Wait ExpEnd

FreeAtom All
FixObj 11
Experiment Minimization
Experiment On
ShowHBoObj 12,Extend=Yes
Wait ExpEnd
```

```
FreeAtom All
FixAtom backbone
FreeObj 12
Experiment Minimization
Experiment On
ShowHBoObj 12,Extend=Yes
Wait ExpEnd

FreeAtom All
Experiment Minimization
Experiment On
ShowHBoObj 12,Extend=Yes
Wait ExpEnd

ShowMessage "Modelling done."
Stop
```

ESI-MS

Table S1: ESI-MS data for compounds **3**-**17**.

Entry	SIM / g·mol^{-1}	$[M+H]^{+}$	$[M+2H]^{2+}$	$[M+3H]^{3+}$	$[M+4H]^{4+}$
3	1647.9	-	824.9	550.3	413.0
4	1619.9	-	811.0	541.0	406.0
5	1691.0	-	846.6	564.7	423.8
6	1719.0	-	860.5	574.0	430.7
7	1658.0	-	830.0	553.7	415.5
8	1605.9	-	804.0	536.3	402.5
9	1665.0	-	833.4	556.0	417.2
10	1663.0	-	832.4	555.3	416.7
11	1416.8	-	709.3	473.2	-
12	1269.6	-	635.8	424.2	-
13	1311.7	-	657.3	438.5	-
14	1374.7	1374.9	688.3	459.2	-
15	1628.9	-	815.5	544.0	408.2
16	1286.6	-	644.2	429.8	-
17	1286.6	-	644.1	429.8	322.6

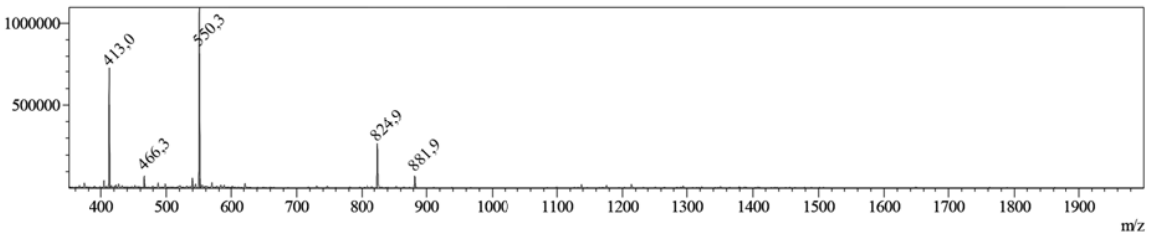


Figure S2. ESI-MS spectrum of compound **3** (positive polarization).

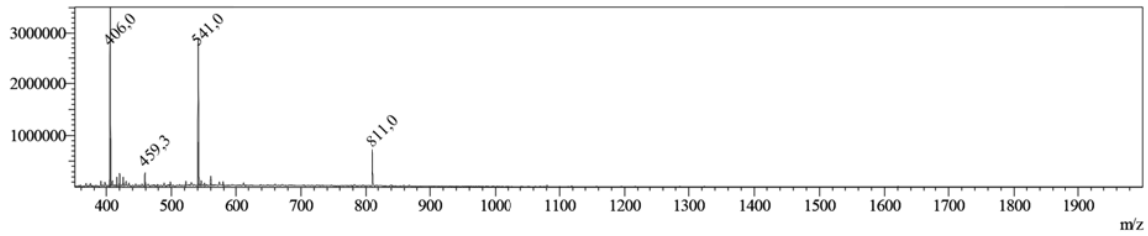


Figure S3. ESI-MS spectrum of compound **4** (positive polarization).

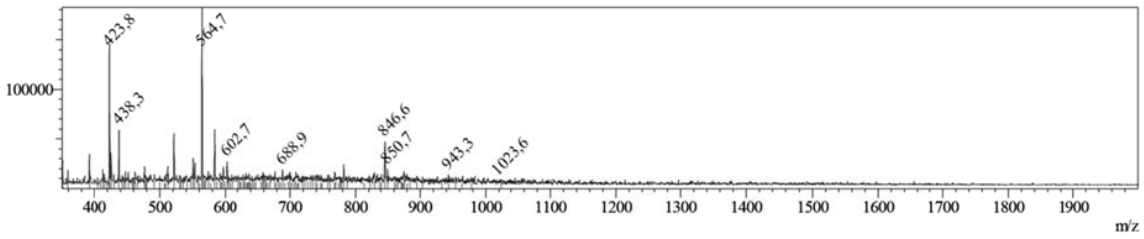


Figure S4. ESI-MS spectrum of compound **5** (positive polarization).

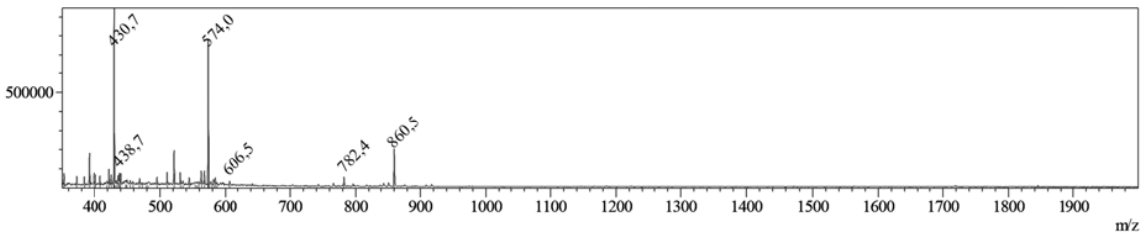


Figure S5. ESI-MS spectrum of compound **6** (positive polarization).

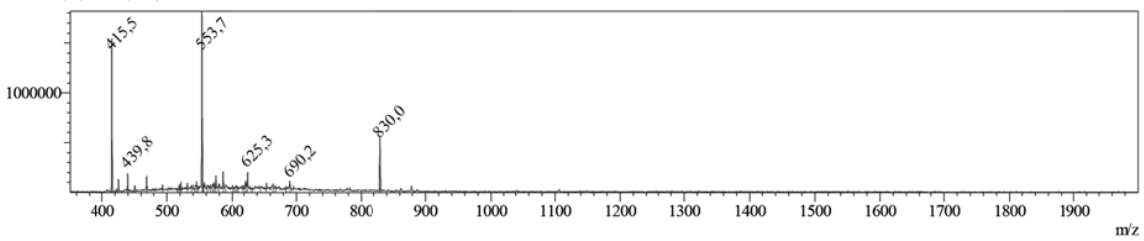


Figure S6. ESI-MS spectrum of compound **7** (positive polarization).

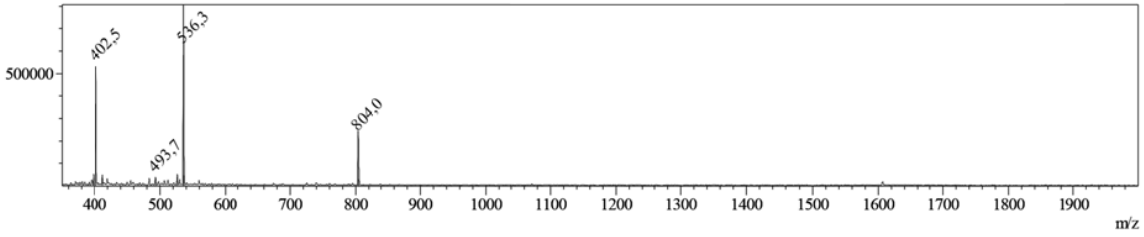


Figure S7. ESI-MS spectrum of compound **8** (positive polarization).

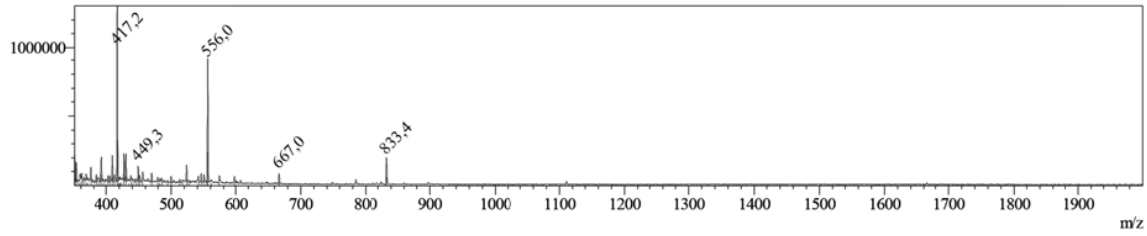


Figure S8. ESI-MS spectrum of compound **9** (positive polarization).

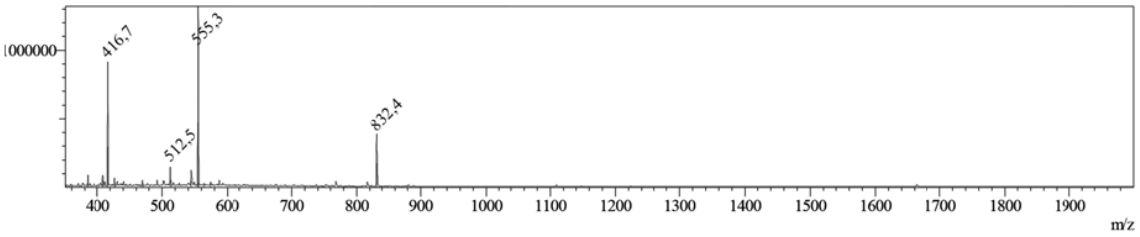


Figure S9. ESI-MS spectrum of compound **10** (positive polarization).

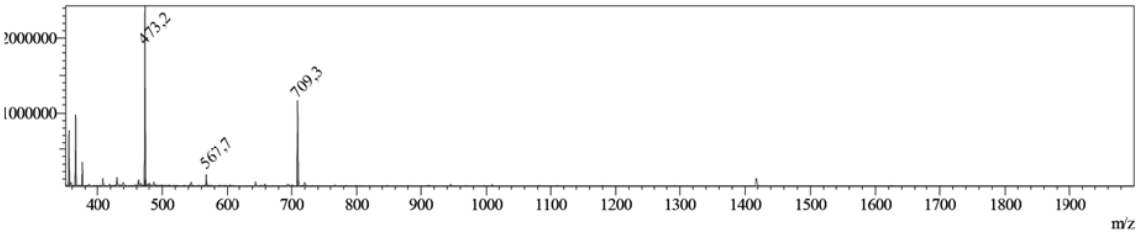


Figure S10. ESI-MS spectrum of compound **11** (positive polarization).

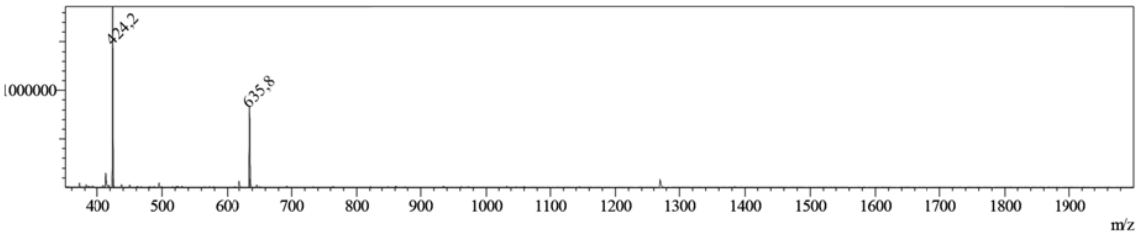


Figure S11. ESI-MS spectrum of compound **12** (positive polarization).

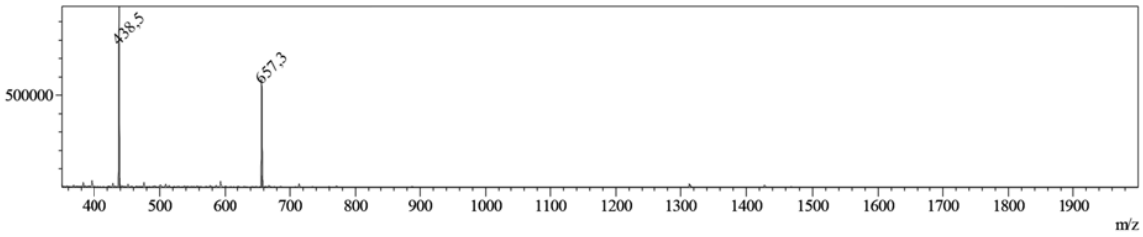


Figure S12. ESI-MS spectrum of compound **13** (positive polarization).

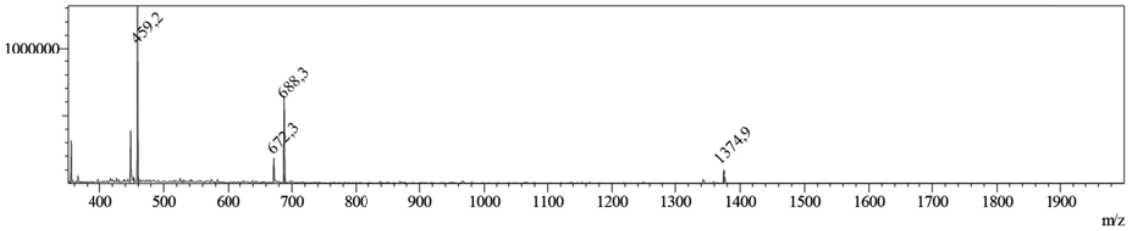


Figure S13. ESI-MS spectrum of compound **14** (positive polarization).

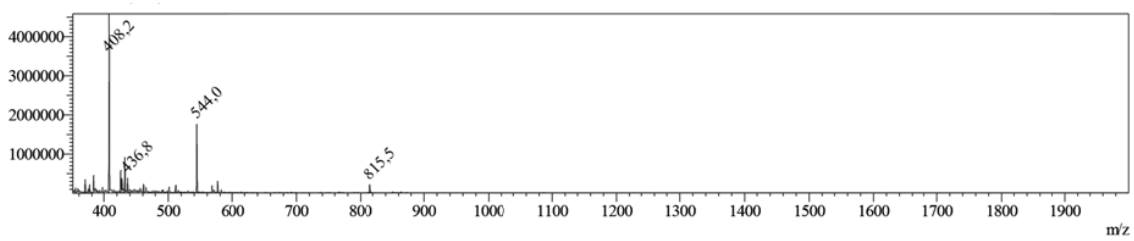


Figure S14. ESI-MS spectrum of compound **15** (positive polarization).

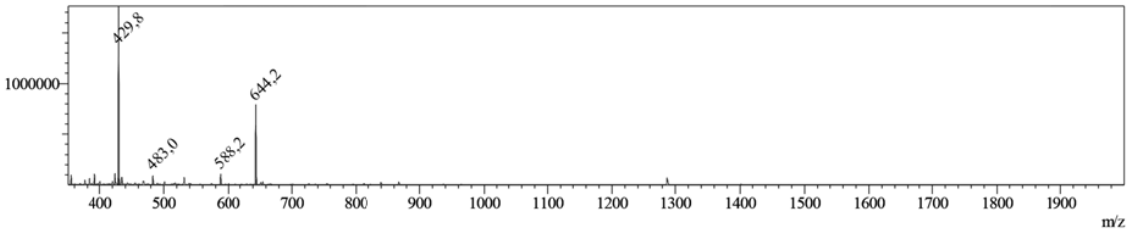


Figure S15. ESI-MS spectrum of compound **16** (positive polarization).

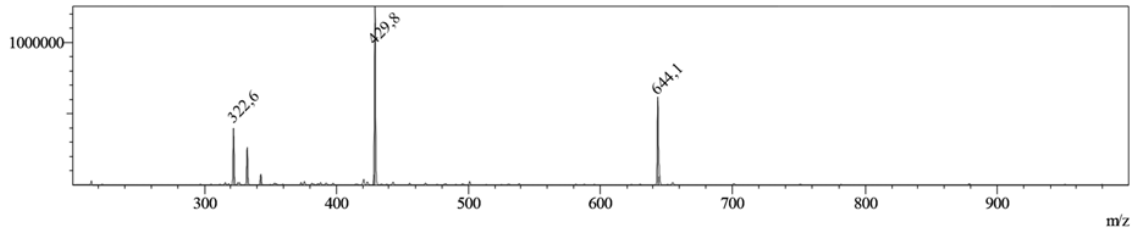


Figure S16. ESI-MS spectrum of compound **17** (positive polarization).

Plotted Kinetic Data

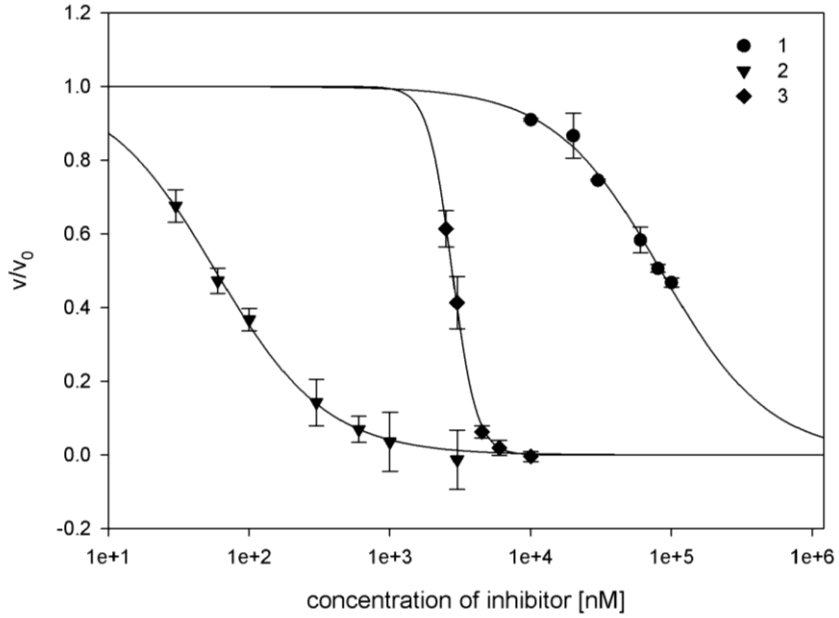


Figure S17. Dose-response curves for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale. Comparison of furin inhibitors **1** (black circle), **2** (black triangle), and **3** (black diamond). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

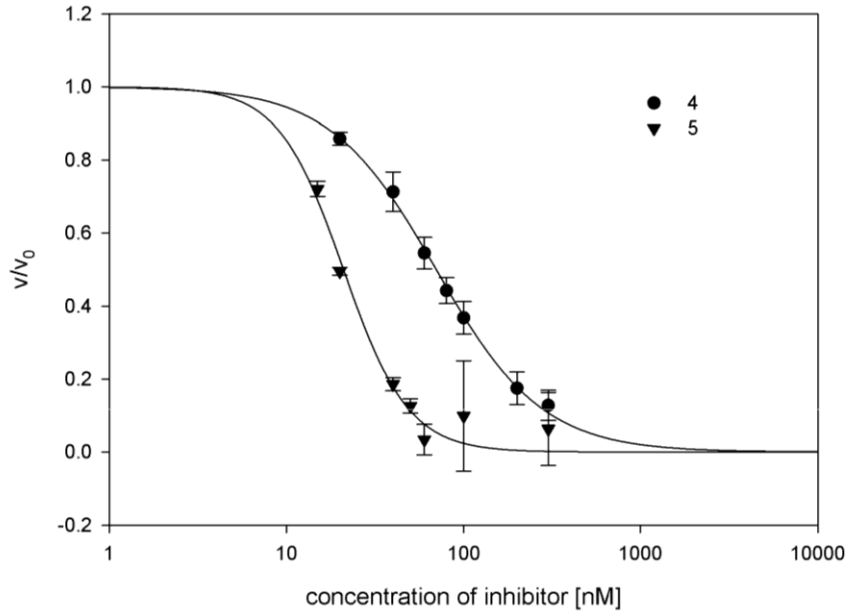


Figure S18. Dose-response curves for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale. Comparison of furin inhibitors **4** (black circle), and **5** (black triangle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

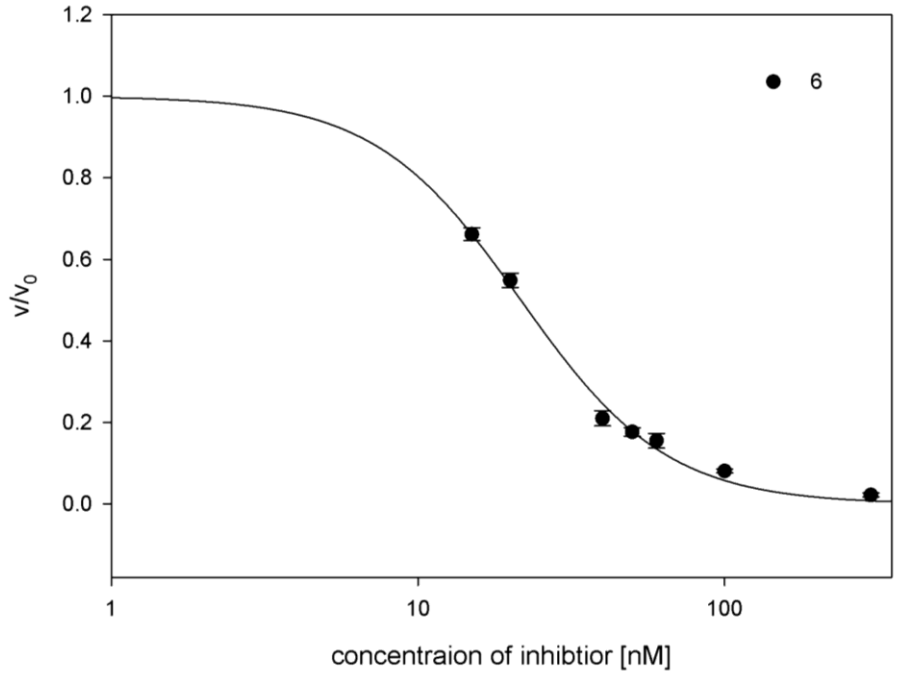


Figure S19. Dose-response curve for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale of compound **6** (black circle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

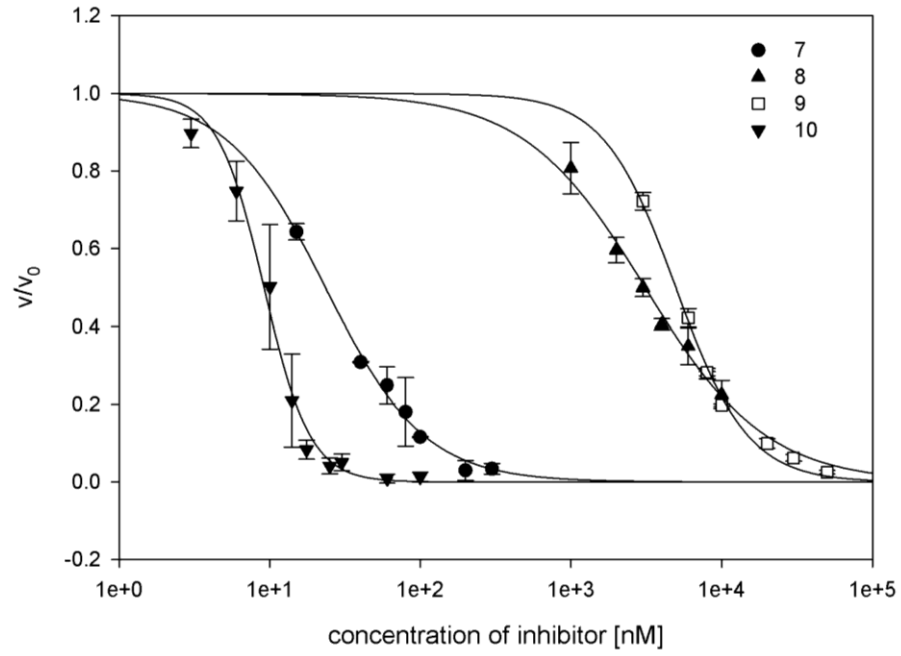


Figure S20. Dose-response curves for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale. Comparison of furin inhibitors **7** (black circle), **8** (black triangle up), **9** (white square), and **10** (black triangle down). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

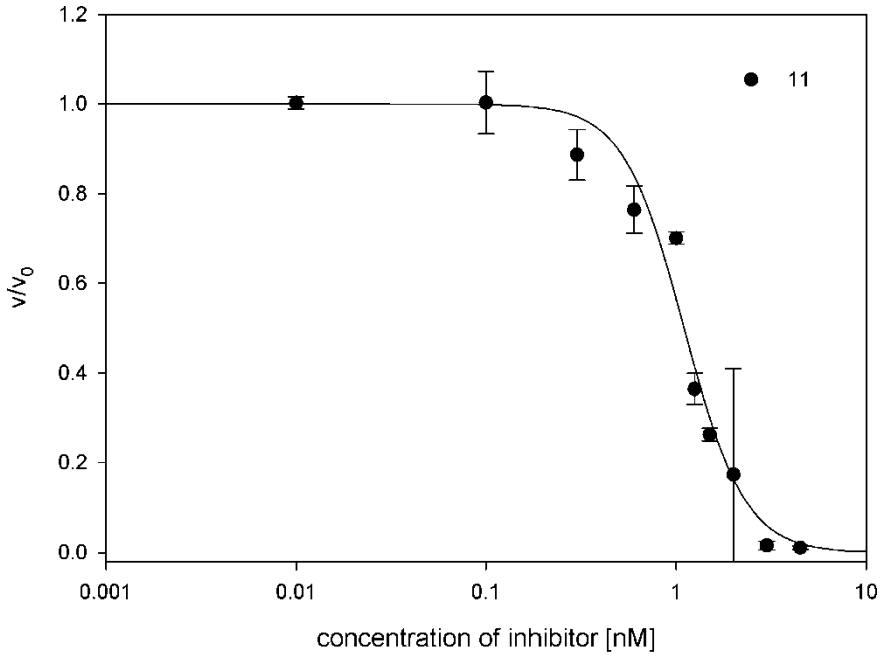


Figure S21. Dose-response curve for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale of compound **11** (black circle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

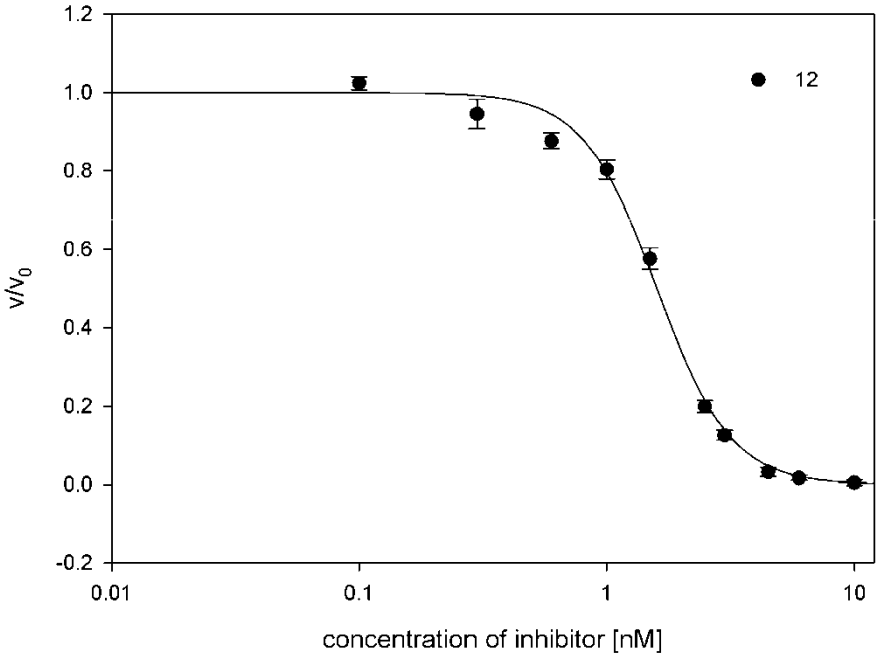


Figure S22. Dose-response curve for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale of compound **12** (black circle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

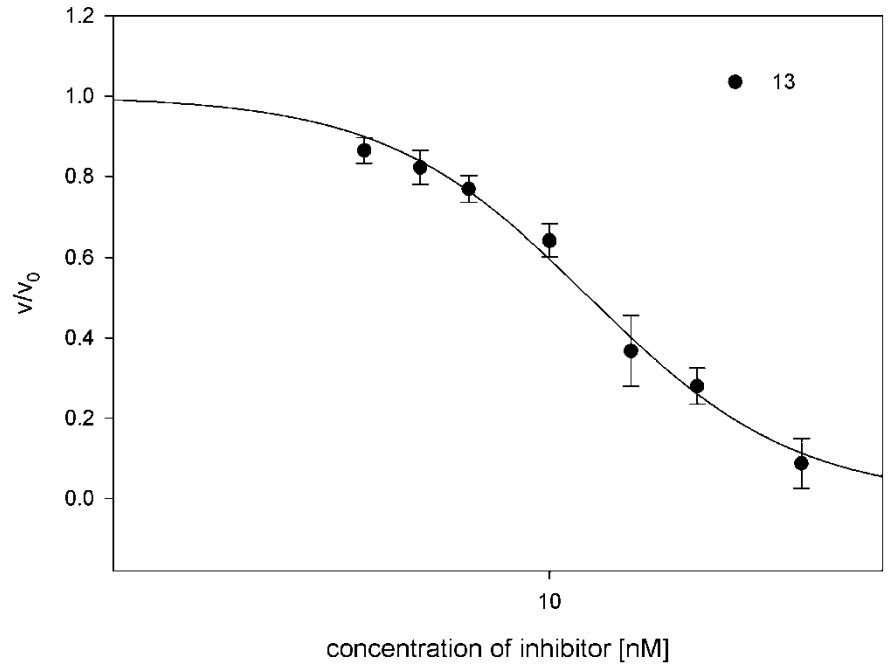


Figure S23. Dose-response curve for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale of compound **13** (black circle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

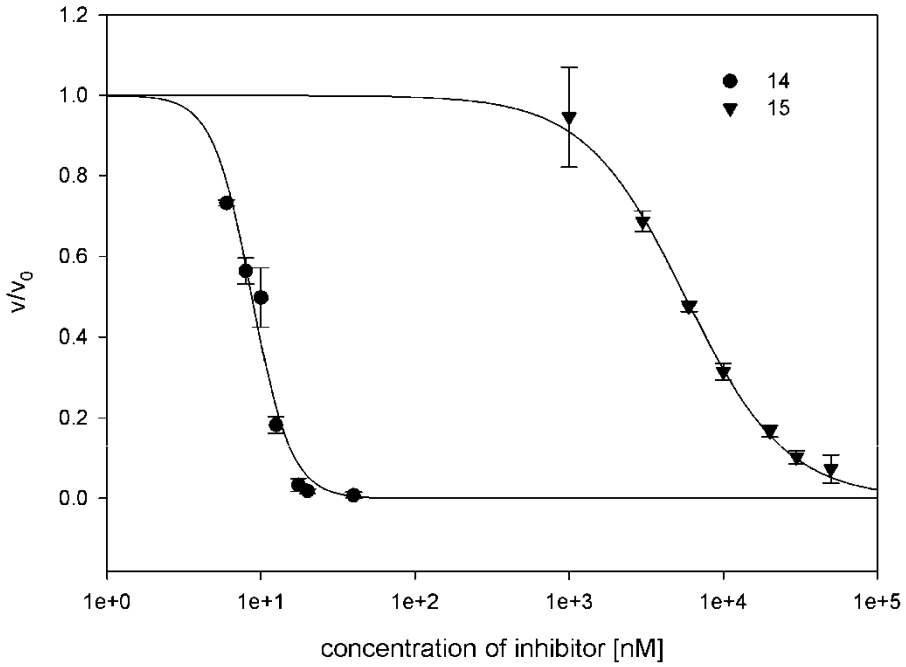


Figure S24. Dose-response curves for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale. Comparison of furin inhibitors **14** (black circle), and **15** (black triangle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

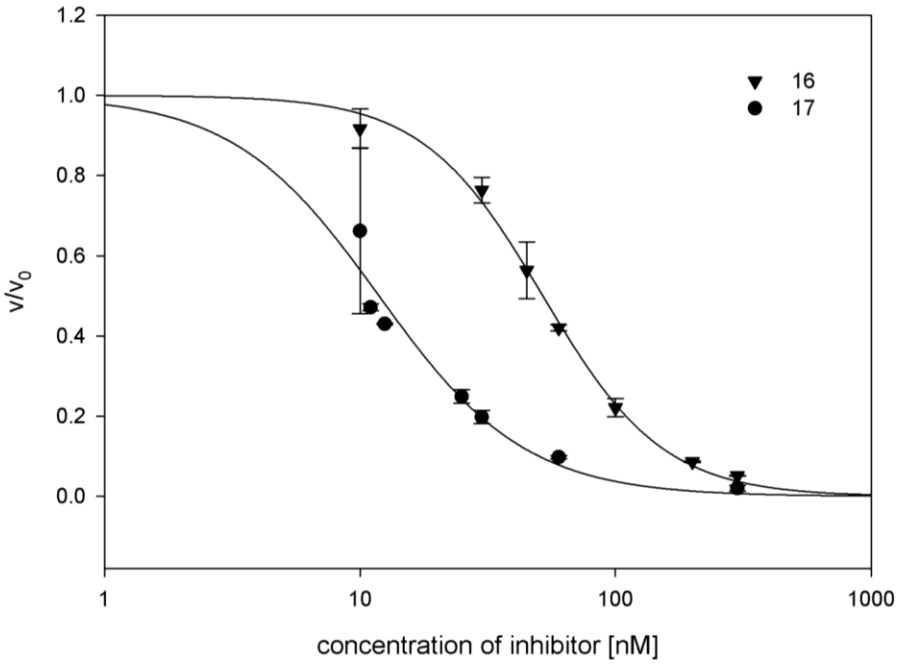


Figure S25. Dose-response curves for the inhibition of furin-catalyzed proteolysis of fluorogenic substrate Boc-RVRR-AMC with the X-axis on a logarithmic scale. Comparison of furin inhibitors **16** (black circle), and **17** (black triangle). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

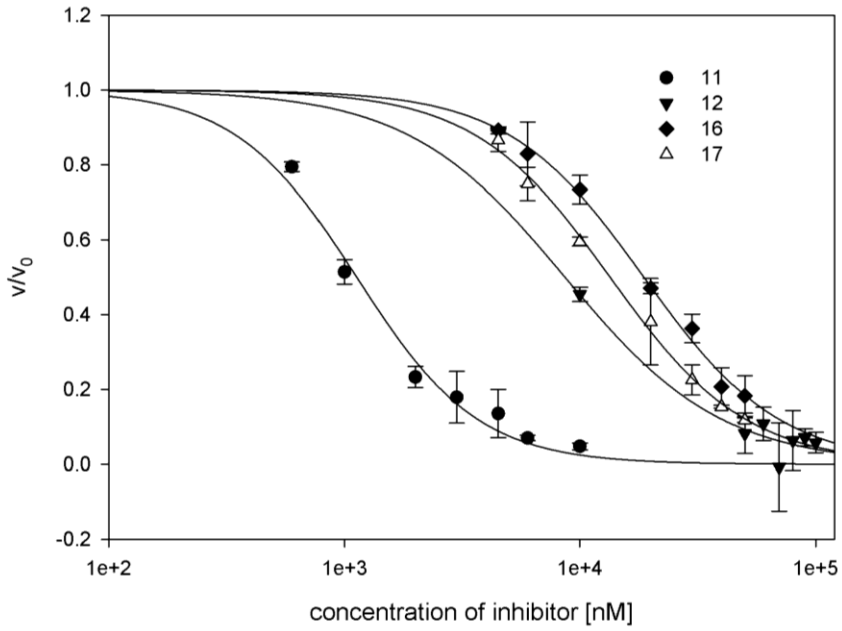


Figure S26. Dose-response curves for the inhibition of matriptase-catalyzed proteolysis of chromophoric substrate Boc-QAR-pNA with the X-axis on a logarithmic scale. Comparison of inhibitors **11** (black circle), **12** (black triangle down), **16** (black diamond), and **17** (white triangle up). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

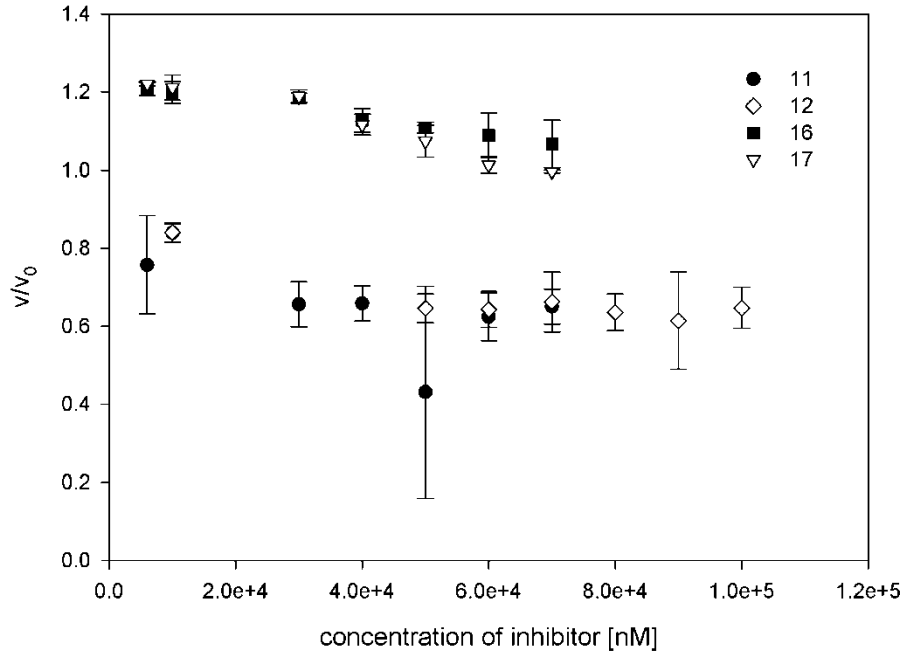


Figure S27. Dose-response curves for the inhibition of trypsin-catalyzed proteolysis of chromophoric substrate Boc-QAR-pNA with the X-axis on a logarithmic scale. Comparison of inhibitors **11** (black circle), **12** (white diamond), **16** (black square), and **17** (white triangle down). Data points are arithmetic means of three experiments and error bars are given as the standard deviation.

HPLC

According to the integration of analytic HPLC traces, the purity of peptides after preparative chromatographic isolation was about 80-97 %, with the lowest values for the Arg-rich peptides (Table S2). Unfortunately, the impurities appeared inherent in the main fraction, and could not be separated even by repeated chromatographic purification. We suppose that this outcome was connected with the quality of used arginine building block.

Table S2: Puriety of the isolated peptides **3-17** *via* RP-HPLC.

Entry	Area [%]
3	80.0
4	84.1
5	85.4
6	81.5
7	92.3
8	85.4
9	90.0
10	84.9
11	93.2
12	87.9
13	80.3
14	81.0
15	83.0
16	92.2
17	97.2

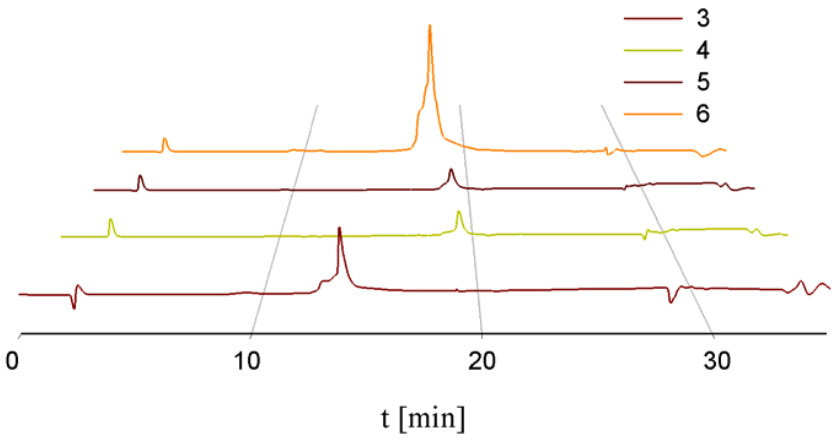


Figure S28. HPLC chromatogram of purified **3-6** recorded at 220 nm.

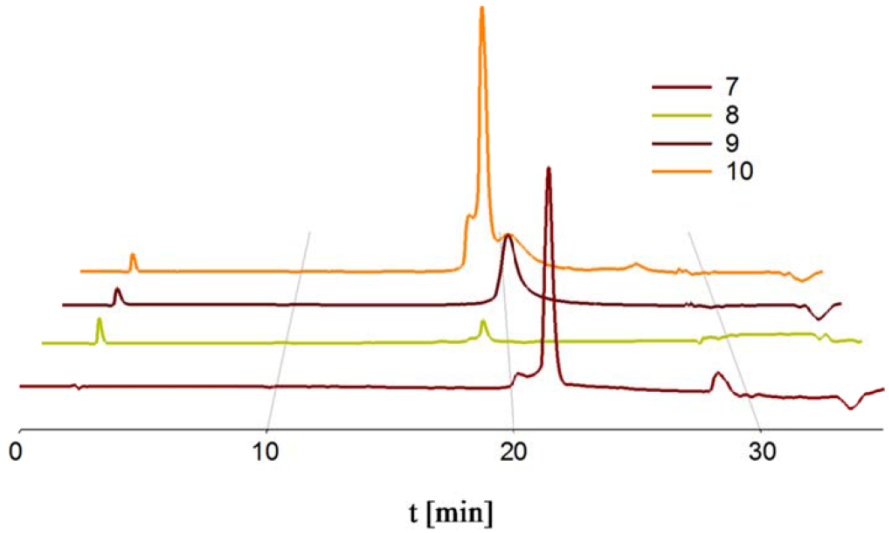


Figure S29. HPLC chromatogram of purified **7**-**10** recorded at 220 nm.

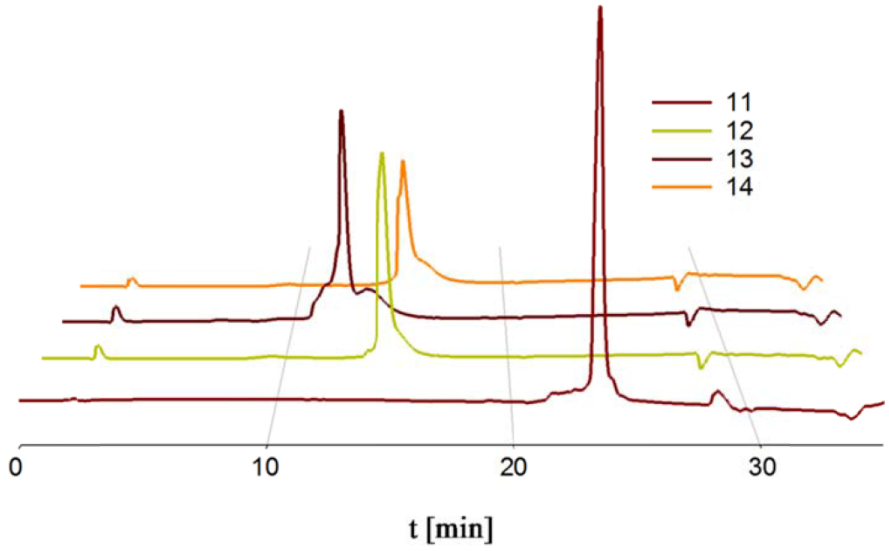


Figure S30. HPLC chromatogram of purified **11**-**14** recorded at 220 nm..

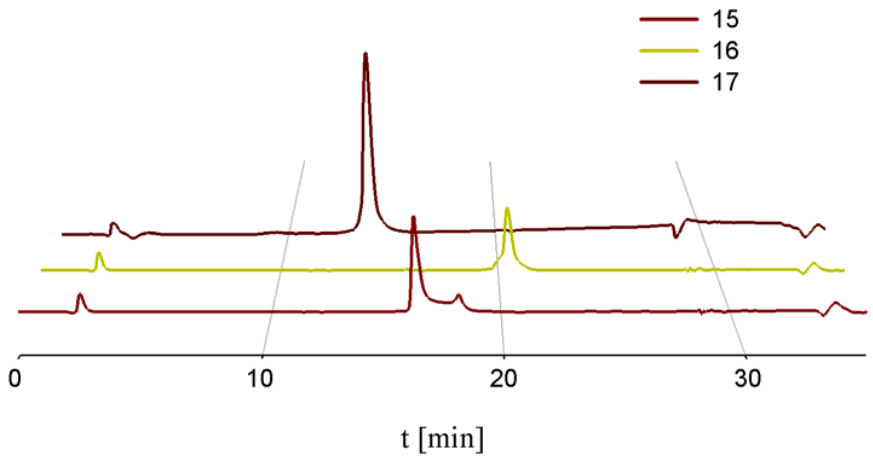


Figure S31. HPLC chromatogram of purified **15-17** recorded at 220 nm.

Literature

[1] O. Avrutina, H. Fittler, B. Glotzbach, H. Kolmar, M. Empting, *Organic & Biomolecular Chemistry* **2012**, *10*, 7753-7762.
[2] S. O. Dahms, K. Hardes, G. L. Becker, T. Steinmetzer, H. Brandstetter, M. E. Than, *ACS Chemical Biology* **2014**, *9*, 1113-1118.
[3] M. Kuester, G. L. Becker, K. Hardes, I. Lindberg, T. Steinmetzer, M. E. Than, *Biological chemistry* **2011**, *392*, 973-981.
[4] E. K. Dufour, A. Désilets, J.-M. Longpré, R. Leduc, *Protein Science* **2005**, *14*, 303-315.
[5] H. Fittler, O. Avrutina, B. Glotzbach, M. Empting, H. Kolmar, *Organic & biomolecular chemistry* **2013**, *11*, 1848-1857.
[6] H. Fittler, O. Avrutina, M. Empting, H. Kolmar, *J Pept Sci* **2014**, *20*, 415-420.
[7] E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker, K. Karplus, *Proteins: Structure, Function, and Bioinformatics* **2009**, *77*, 114-122.

7. Appendix

List of Figures

Figure 1: Secondary structure of human matriptase-1 (PDB: 3P8F) (left) with a close-up of the three amino acids at the active site (right).....4
Figure 2: Depiction of the oxyanion hole of chymotrypsin. The tetrahedral intermediate is stabilized by the amine groups of Gly^{193} and Ser^{195}. 4
Figure 3: Catalytic mechanism of a serine protease. The carbonyl carbon of the substrate is attacked by the hydroxyl of serine, followed by the cleavage of the amine. A water molecule adds to the carbonyl carbon and the cycle can be repeated. 5
Figure 4: Left: Crystal structure of matriptase with elements of secondary structure indicated (PDB: 3P8F). Right: Multidomain structure of matriptase-1 comprising the *N*-terminal cytoplasmic domain and membrane, the extracellular domain containing the stem region (including SEA and two CUB domains) and four LDLRA (low-density lipoprotein receptor domain) domains, as well as the catalytic triad at the *C*-terminus (figure modified from K. Uhland).[45] 6
Figure 5: Selected potent matriptase-1 inhibitors.[52-55, 59, 62] 7
Figure 6: Left: Crystal structure of furin with elements of secondary structure indicated (PDB: 1P8J). Right: Sketch of multidomain structure of furin with its catalytic triad shown in one-letter amino acid code.[64, 66-67] 7
Figure 7: Selected potent furin inhibitors.[74-76] 8
Figure 8: Comparison of the molecular weight of an antibody (PDB: 1IGT), EETI (PDB: 2IT7), and SFTI-1 (PDB: 1SFI).....9
Figure 9: Depiction of different disulfide-rich miniproteins. a) MCoTI-II (PDB: 1HA9), b) EETI (PDB: 2IT7) c), Kalata B1 (PDB: 2LUR) d) AgRP (PDB: 1MR0).[80] 9
Figure 10: Active site sequence of Bowman-Birk inhibitors in comparison. A conserved disulfide bond formed by two cysteines (bold), and a *cis*-Pro at the P3' position (bold) are common in all BBIs. Origin and length of the molecules are also stated (figure modified from Daly *et al.*).[94] 10
Figure 11: Depiction of the sunflower trypsin inhibitor-1 (SFTI-1) (PDB: 1SFI). Blue: nitrogen, red: oxygen, yellow: sulfur, grey: carbon, hydrogen atoms are omitted for clarity.....11

Lebenslauf

M. Sc.
Heiko Fittler

Wienerstraße 91
64287 Darmstadt
E-Mail: fittler@biochemie-tud.de

Persönliche Daten

Geburtsdatum	13. Oktober 1984
Geburtsort	Aschaffenburg
Staatsangehörigkeit	deutsch

Ausbildung und Berufserfahrung

04/2012-08/2015	**Wissenschaftlicher Mitarbeiter der TU Darmstadt** Promotion in der Chemie (Fachgebiet Biochemie) **Projekt**: Synthese von neuartigen Peptidinhibitoren gegen tumorrelevante Ziele (Matriptase, Furin)
10/2009-03/2012	**Master of Science (M. Sc.) in Chemie** Technische Universität Darmstadt **Thesis**: Scaffold-based modular approach towards novel protease inhibitors
10/2006-09/2009	**Bachelor of Science (B. Sc.) in Chemie** Technische Universität Darmstadt **Thesis**: Design and synthesis of bioactive peptides and peptide multimers using conjugation templates and miniprotein scaffolds
07/2006-10/2006	**TUJA** Zeitarbeitsfirma
10/2005-06/2006	**Zivildienst** Bayrisches Rotes Kreuz (BRK)
08/2002-07/2005	**Gymnasium – Allgemeine Hochschulreife** Ludwig-Geissler-Schule in Hanau
08/1995-07/2002	**Realschule - Fachoberschulreife** Realschule für Knaben in Aschaffenburg

M. Sc.
Heiko Fittler
Wienerstraße 91
64287 Darmstadt
E-Mail: fittler@biochemie-tud.de

Darmstadt, den

Eidessattliche Erklärung

Ich erkläre hiermit, dass ich meine Dissertation selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

_____

M. Sc. Heiko Fittler

M. Sc.
Heiko Fittler
Wienerstraße 91
64287 Darmstadt
E-Mail: fittler@biochemie-tud.de

Darmstadt, den

Erklärung

Ich erkläre hiermit, noch keinen Promotionsversuch unternommen zu habe.

_____

M. Sc. Heiko Fittler